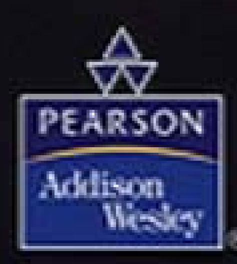
TERCERA EDICIÓN

Álgebra Lineal

Y SUS APLICACIONES

Actualizada



David C. Lay

Álgebra lineal

y sus aplicaciones

Álgebra lineal
y sus aplicaciones

TERCERA EDICIÓN ACTUALIZADA

David C. Lay
University of Maryland – College Park

TRADUCCIÓN
Jesús Elmer Murrieta Murrieta

México • Argentina • Brasil • Colombia • Costa Rica • Chile • Ecuador
España • Guatemala • Panamá • Perú • Puerto Rico • Uruguay • Venezuela

Datos de catalogación bibliográfica

LAY, DAVID C.

ÁLGEBRA LINEAL Y SUS APLICACIONES

PEARSON EDUCACIÓN, México, 2007
ISBN: 978-970-26-0906-3
Área: Matemáticas

Formato: 20 × 25.5 cm Páginas: 584

ISBN 0321287134

Edición en español:
Editor: Luis Miguel Cruz Castillo
e-mail: luis.cruz@pearsoned.com
Editor de desarrollo: Claudia Martínez Amigon
Supervisor de producción: Adriana Rida Montes

Edición en inglés:
Publisher: Greg Tobin
Acquisitions Editor: William Hoffman
Project Editor: Joanne Ha
Editorial Assistant: Emily Portwood
Managing Editor: Karen Wernholm
Production Supervisor: Sheila Spinney
Senior Designer/Cover Designer: Barbara T. Atkinson
Photo Researcher: Beth Anderson
Digital Assets Manager: Jason Miranda
Media Producer: Sara Anderson
Software Development: David Malone y Mary Durnwald
Marketing Manager: Phyllis Hubbard
Marketing Coordinator: Celena Carr
Senior Author Support/Technology Specialist: Joe Vetere
Rights and Permissions Advisor: Dana Weightman
Senior Manufacturing Buyer: Evelyn Beaton
Composition: Techsetters, Inc.
Illustrations: Techsetters, Inc.
Photo Credits: 1 Bettmann/Corbis; Hulton Archive. 58, 63, 98, 156, 185, 252, 426, 469 PhotoDisc. 105 The Boeing Company. 106 Boeing Phantom Works. 140 Jet Propulsion Lab/NASA. 161 Bo Strain; Reprinted by permission of University of North Carolina at Chapel Hill. 215 Kennedy Space Center. 289, 469 Eyewire. 301 Stone. 373 Corbis. 374 From North American Datum of 1983, Charles Schwartz editor, National Geodetic Information Center. 426 Anglo-Australian Observatory/Royal Observatory, Edinburgh. 447 NASA. 448 GEOPIC image courtesy of Earth Satellite Corporation, Rockville, MD.

TERCERA EDICIÓN, 2007

Cámara Nacional de la Industria Editorial Mexicana.

Reg. Núm. 1031.

ISBN 10: 970-26-0906-2
ISBN 13: 978-970-26-0906-3

Impreso en México. *Printed in Mexico.*

1 2 3 4 5 6 7 8 9 0 - 10 09 08 07

A mi esposa, Lillian, y a nuestras hijas
Christina, Deborah y Melissa,
cuyo apoyo, ánimos, y fieles oraciones
hicieron posible este libro

Acerca del autor

David C. Lay tiene los títulos de B. A. de Aurora University (Illinois), y de M. A. y PH. D. por la Universidad de California en Los Ángeles. El profesor Lay ha sido catedrático e investigador en matemáticas desde 1966, principalmente en la Universidad de Maryland, College Park. También ha trabajado como profesor visitante en la Universidad de Ámsterdam, en la Universidad Libre de Ámsterdam y en la Universidad de Kaiserslautern, Alemania. Tiene más de treinta artículos de investigación publicados como análisis funcional y álgebra lineal.

Como miembro fundador del Grupo de Estudio del Currículum de Álgebra Lineal patrocinado por la N.S.F., el profesor Lay ha sido líder en el movimiento actual para modernizar el plan de estudios de álgebra lineal. El profesor Lay también es coautor de varios textos matemáticos, entre ellos, *Introduction to Functional Analysis*, con Angus E. Taylor, *Calculus and its Applications*, con L. J. Goldstein y D. I. Schneider, y *Linear Algebra Gens – Assets for Undergraduate Mathematics*, con D. Carlson, C. R. Johnson y A. D. Porter.

Catedrático de primera línea. El profesor Lay ha recibido cuatro premios universitarios por excelencia docente, incluido en 1996 el de Distinguished Scholar–Teacher de la Universidad de Maryland. En 1994, se le concedió uno de los Premios de la Mathematical Association of America, que lleva el título de Distinguished College or University Teaching of Mathematics. Ha sido elegido por los estudiantes universitarios miembro de la Alpha Lambda Delta National Scholastic Honor Society y de la Golden Key National Honor Society. En 1989, la Aurora University le concedió el premio Outstanding Alumnus. El doctor Lay es miembro de la American Mathematical Society, de la Canadian Mathematical Society, de la International Linear Algebra Society, de la Mathematical Association of America, Sigma Xi, y de la Society for Industrial and Applied Mathematics. Desde 1992, ha formado parte de la junta directiva nacional de la Association of Christians in the Mathematical Sciences.

Contenido

Prefacio

La respuesta de estudiantes y profesores a las primeras tres ediciones de *Álgebra lineal y sus aplicaciones* ha sido muy gratificante. Esta *tercera edición actualizada* proporciona un apoyo sustancial tanto para la enseñanza como para el uso de tecnología en el curso. Como antes, el texto presenta una introducción elemental moderna al álgebra lineal y una amplia selección de interesantes aplicaciones. El material es accesible a estudiantes que hayan adquirido la madurez necesaria, por lo general, en cálculo, después de completar satisfactoriamente dos semestres de matemáticas a nivel universitario.

La meta principal del texto es ayudar a los estudiantes a dominar los conceptos y las habilidades básicas que después utilizarán en sus carreras. Los temas incluidos siguen las recomendaciones del Linear Algebra Curriculum Study Group, las cuales se basan en una investigación cuidadosa de las necesidades reales de los estudiantes y en un consenso logrado entre profesionales de muchas disciplinas que utilizan álgebra lineal. Espero que este curso sea una de las clases de matemáticas más útiles e interesantes que puedan tomarse durante los estudios universitarios.

CARACTERÍSTICAS DISTINTIVAS

Introducción temprana de conceptos clave

Muchas ideas fundamentales del álgebra lineal se introducen en siete lecturas, una lectura al inicio de cada capítulo, en el establecimiento concreto de $\mathbb{R}^n$, y después se examinan de manera gradual desde diferentes puntos de vista. Posteriormente aparecen generalizaciones de estos conceptos como extensiones naturales de ideas familiares, visualizadas a través de la intuición geométrica desarrollada en el capítulo 1. En la opinión del autor, una de las características positivas del texto es que el nivel de dificultad es bastante uniforme a lo largo del curso.

Una visión moderna de la multiplicación de matrices

La notación correcta es crucial, y el texto refleja la forma real en que los científicos e ingenieros aplican el álgebra lineal en la práctica. Las definiciones y comprobaciones se enfocan en las columnas de una matriz en lugar de en sus entradas. Un tema esencial es considerar un producto vector-matriz $A\mathbf{x}$ como una combinación lineal de las columnas de A. Este moderno enfoque simplifica muchos argumentos, y vincula las ideas de espacio vectorial con el estudio de sistemas lineales.

Transformaciones lineales

Las transformaciones lineales forman un “hilo” que se entreteje en la tela de este texto. Su utilización mejora el sentido geométrico de lo escrito. Por ejemplo, en el capítulo 1, las transformaciones lineales proporcionan una visión dinámica y gráfica de la multiplicación matriz-vector.

Valores propios y sistemas dinámicos

Los valores propios aparecen equitativamente pronto en el texto, en los capítulos 5 y 7. Como este material se estudia durante varias semanas, los alumnos tienen más tiempo del usual para absorber y revisar estos conceptos críticos. Los valores propios se aplican a sistemas dinámicos discretos y continuos, los cuales aparecen en las secciones 1.10, 4.8, 4.9, y en cinco secciones del capítulo 5. Algunos cursos llegan al capítulo 5 en unas cinco semanas pues cubren las secciones 2.8 y 2.9 en lugar del capítulo 4. Estas dos secciones opcionales presentan todos los conceptos del espacio vectorial incluidos en el capítulo 4, mismos que son necesarios para abordar el capítulo 5.

Ortogonalidad y problemas de mínimos cuadrados

Estos temas reciben un tratamiento más comprensible en comparación con el que se encuentra comúnmente en los textos básicos. El Linear Algebra Curriculum Study Group ha enfatizado la necesidad de contar con una unidad sustancial en los problemas de ortogonalidad y mínimos cuadrados, debido a que la ortogonalidad cumple un papel importante en los cálculos computacionales y en el álgebra lineal numérica, y porque los sistemas lineales inconsistentes surgen muy frecuentemente en el trabajo práctico.

CARACTERÍSTICAS PEDAGÓGICAS

Aplicaciones

Una amplia selección de aplicaciones ilustra el poder del álgebra lineal para explicar principios fundamentales y simplificar los cálculos en ingeniería, ciencia computacional, matemáticas, física, biología, economía y estadística. Algunas aplicaciones aparecen en secciones diferentes; otras se explican mediante ejemplos y ejercicios. Además, cada capítulo abre con un ejemplo introductorio que especifica la etapa apropiada para efectuar determinada aplicación del álgebra lineal, y proporciona una motivación para desarrollar las matemáticas que siguen. Después, el texto retoma la aplicación en una sección cercana al final del capítulo.

Un fuerte énfasis geométrico

En el curso, todos los conceptos importantes reciben una interpretación geométrica, debido a que muchos estudiantes aprenden de mejor manera cuando pueden visualizar una idea. Existe una cantidad sustancialmente mayor de ilustraciones de lo usual, y algunas de las figuras no han aparecido nunca antes en un texto de álgebra lineal.

Ejemplos

En contraste con lo que se acostumbra en la mayor parte de los libros de álgebra, este texto dedica una proporción más grande de su material de exposición a ejemplos. Existen más ejemplos de los que ordinariamente presentaría un profesor en clase. Pero como han sido escritos con cuidado y de manera detallada, los estudiantes pueden leerlos por sí mismos.

Teoremas y demostraciones

Los resultados importantes se establecen como teoremas. Otros conceptos útiles se despliegan dentro de recuadros iluminados para utilizarse como referencias rápidas. La mayor parte de los teoremas tienen comprobaciones formales, escritas pensando en los alumnos principiantes. En algunos casos, los cálculos esenciales de una comprobación se muestran en un ejemplo seleccionado cuidadosamente. Algunas verificaciones de rutina se dejan para la sección de ejercicios, cuando esto resulta benéfico para los estudiantes.

Problemas de práctica

Antes de cada serie de ejercicios aparecen algunos problemas de práctica seleccionados en forma cuidadosa. La serie de ejercicios va seguida por soluciones completas. Estos problemas se enfocan en dificultades potenciales que pueden encontrarse en la serie de ejercicios o proporcionan un "calentamiento" para la ejecución posterior de los ejercicios; con frecuencia, las soluciones contienen sugerencias o advertencias útiles acerca de la tarea.

Ejercicios

La abundancia de ejercicios incluye desde cálculos de rutina hasta preguntas conceptuales que requieren de mayor reflexión. Un buen número de preguntas innovadoras destacan las dificultades conceptuales que el autor ha encontrado en los estudiantes a través de los años. Cada serie de ejercicios se organiza cuidadosamente, en el mismo orden general que el texto: las asignaciones de tarea pueden encontrarse con facilidad cuando sólo se ha estudiado una parte de determinada sección. Una característica notable de los ejercicios es su simplicidad numérica. Los problemas se "desdoblan" rápidamente, por lo que los estudiantes pasan poco tiempo realizando cálculos numéricos. Los ejercicios se concentran en inducir la comprensión de los temas, en vez de demandar cálculos mecánicos.

Preguntas de verdadero o falso

Para estimular a los estudiantes a leer todo el texto y a pensar de manera crítica, se han desarrollado 300 preguntas simples del tipo verdadero o falso que aparecen en 33 secciones del texto, justo enseguida de los problemas computacionales. Estas preguntas pueden responderse directamente a partir del texto y preparan al estudiante para los problemas conceptuales que vienen después. Los estudiantes aprecian estas preguntas —luego de reconocer la importancia de leer el texto con cuidado—. Con base en pruebas de clase y discusiones con estudiantes, se decidió no poner las respuestas en el texto. Para comprobar la comprensión del material, existen 150 preguntas adicionales del tipo verdadero o falso (casi siempre al final de los capítulos.) El texto proporciona respuestas simples V/F a la mayor parte de estas preguntas, pero omite las justificaciones a las respuestas (que, por lo general, requieren de cierta reflexión).

Ejercicios de escritura

Para todos los estudiantes de álgebra lineal resulta esencial poseer la capacidad de escribir enunciados matemáticos coherentes, no sólo para quienes obtendrán un título en matemáticas. El texto incluye muchos ejercicios para los cuales parte de la respuesta consiste en proporcionar una justificación escrita. Los ejercicios conceptuales que requieren una comprobación corta contienen, por lo general, sugerencias que ayudan al estudiante a iniciar la búsqueda de la solución. Para gran parte de los ejercicios de escritura con número impar, se incluye una solución al final del texto o se proporciona una sugerencia.

Temas computacionales

El texto acusa el impacto de la computadora tanto en el desarrollo como en la práctica del álgebra lineal en las ciencias y la ingeniería. Las frecuentes notas numeradas dirigen la atención hacia aspectos de cómputo y distinguen entre conceptos teóricos, digamos la inversión de matrices, e implementaciones de computadora, tales como las factorizaciones LU.

CD ANEXO Y SOPORTE EN LA RED

La edición actualizada del texto incluye una copia completa (en inglés) de la *Guía de estudio* (S*tudy Guide*) en el CD anexo. Esta *guía* fue escrita para ser una parte integral del curso. Un ícono SG en el texto dirige a los estudiantes a subsecciones especiales de la *guía* que sugieren cómo dominar los conceptos clave del curso. La *guía* proporciona una solución detallada a cada tercer ejercicio con número impar, lo que permite a los estudiantes verificar su trabajo. Se proporciona una explicación completa cada vez que un ejercicio de escritura con número impar tiene sólo una “sugerencia” en las respuestas. Existen “advertencias” frecuentes que identifican los errores comunes y muestran cómo evitarlos. Los recuadros de MATLAB presentan comandos cada vez que uno de éstos es necesario. Los apéndices en la *Guía de estudio* proporcionan información comparable acerca de Maple, Mathematica y calculadoras gráficas TI y HP.

Inicio del trabajo con tecnología

Si su curso incluye algún trabajo con MATLAB, Maple, Mathematica o calculadoras TI o HP, puede leer uno de los proyectos que aquí se presentan para obtener una introducción a la tecnología. (Vea la página 104 del texto.)

Archivos de datos

Cientos de archivos contienen datos para alrededor de 900 ejercicios numéricos incluidos en el texto, estudios de caso y proyectos de aplicación. Los datos están disponibles en una diversidad de formatos —para MATLAB, Maple, Mathematica y las calculadoras gráficas TI-83+/86/89 y HP48G. Al permitir a los estudiantes la introducción de matrices y vectores para un problema en particular con unos cuantos golpes de tecla, los archivos de datos eliminan errores de entrada y ahorran tiempo en la realización de tareas.

Nuevos proyectos de MATLAB

Estos proyectos exploratorios invitan a los estudiantes a descubrir aspectos matemáticos y numéricos que son básicos en álgebra lineal. Escritos por Rick Smith, fueron desarrollados para acompañar un curso computacional de álgebra lineal en University of Florida, donde se ha utilizado *Álgebra lineal y sus aplicaciones* por muchos años. Los proyectos están señalados mediante el ícono CD en puntos adecuados del texto. Alrededor de la mitad de los proyectos exploran conceptos fundamentales como el espacio de columna, la diagonalización, y las proyecciones ortogonales; otros se enfocan en aspectos numéricos como los *flops*, métodos iterativos, y la DVS, y algunos examinan aplicaciones como las cadenas de Markov.

www.pearsoneducacion.net/lay

Esta página web contiene el material incluido en el CD anexo, excepto la *Guía de estudio* y los nuevos proyectos de MATLAB. Además, el sitio contiene el primer capítulo

del texto actualizado y el primer capítulo de la *Guía de estudio* (en inglés). Este material es proporcionado para ayudar a los profesores a iniciar con su curso, tal como si una librería distribuyera el texto justo antes de que las clases comenzaran. Para los estudiantes, el sitio en red contiene **hojas de repaso** y **exámenes de práctica** (con soluciones) que cubren los temas principales del texto. Provienen de manera directa de cursos que el autor ha impartido en los últimos años. Cada hoja de repaso identifica definiciones clave, teoremas y habilidades de una parte específica del texto.

Aplicaciones por capítulos

El sitio en la red también contiene siete casos de estudio, los cuales amplían los temas introducidos al inicio de cada capítulo al agregar datos del mundo real y oportunidades para efectuar una exploración más profunda. Por otro lado, más de veinte proyectos de aplicación hacen extensivos los temas del texto o introducen nuevas aplicaciones, como ranuras cúbicas, rutas de vuelo en aerolíneas, matrices de dominancia en competencias deportivas, y códigos de corrección de errores. Algunas aplicaciones matemáticas son las técnicas de integración, la localización de raíces polinomiales, las secciones cónicas, las superficies cuadráticas, y los extremos para funciones de dos variables. También se incluyen temas de álgebra lineal numérica, como números de condición, factorización de matrices, y el método QR para encontrar valores propios. Entrelazados en cada análisis se encuentran ejercicios que pueden involucrar grandes series de datos (y por ende requerir el uso de la tecnología para resolverlos).

RECURSOS PARA EL PROFESOR

Página de recursos para profesores

En la página Web www.pearsoneducacion.net/lay el profesor también puede acceder a una página de descarga donde encontrará todos los archivos de los materiales que acompañan al libro de texto. Entre otras cosas, esta página incluye:

- Manual de soluciones a los ejercicios del libro.
- Banco de exámenes en formato electrónico.
- Dos capítulos adicionales a los del libro impreso.
- Manuales de las aplicaciones y calculadoras más utilizadas.

Curso de CourseCompass en línea

Este libro cuenta también con un curso precargado en CourseCompass, que es una plataforma completa para cursos en línea desarrollada por Blackboard Technologies y complementada con contenidos de Pearson Educación. En ésta el profesor puede asignar exámenes y tareas, organizar todos los materiales del curso, comunicarse con sus alumnos y administrar las calificaciones. Para mayor información, visite www.pearsoneducacion.net/coursecompass

RECONOCIMIENTOS

El autor expresa su gratitud a muchos grupos de personas que lo han ayudado a través de los años con diferentes aspectos del libro.

Se agradece a Israel Gohberg y Robert Ellis por más de quince años de colaboración en la investigación del álgebra lineal, lo cual ha conformado en gran medida una visión particular de esta materia.

Ha sido un privilegio trabajar con David Carlson, Charles Johnson, y Duane Porter en el *Linear Algebra Curriculum Study Group*. Sus ideas sobre la enseñanza del álgebra lineal han influido en este texto de muchas maneras importantes.

Agradezco de manera sincera a los siguientes revisores por su análisis cuidadoso y sus sugerencias constructivas:

Revisores de la tercera edición y ejecutores de pruebas en clase

David Austin, *Grand Valley State University*
G. Barbanson, *University of Texas at Austin*
Kenneth Brown, *Cornell University*
David Carlson, *San Diego State University*
Greg Conner, *Brigham Young University*
Casey T. Cremins, *University of Maryland*
Sylvie DesJardins, *Okanagan University College*
Daniel Flath, *University of South Alabama*
Yuval Flicker, *Ohio State Universitv*
Scott Fulton, *Clarkson University*
Herman Gollwitzer, *Drexel University*
Jeremy Haefner, *University of Colorado at Colorado Springs*
William Hager, *University of Florida*
John Hagood, *Northern Arizona University*
Willy Hereman, *Colorado School of Mines*
Alexander Hulpke, *Colorado State University*
Doug Hundley, *Whitman College*
James F. Hurley, *University of Connecticut*
Jurgen Hurrelbrink, *Louisiana State University*
Jerry G. Ianni, *La Guardia Community College (CUNY)*
Hank Kuiper, *Arizona State University*
Ashok Kumar, *Valdosta State University*
Earl Kymala, *California State University*, *Sacramento*
Kathryn Lenz, *University of Minnesota-Duluth*
Jaques Lewin, *Syracuse University*
En-Bing Lin, *University of Toledo*
Andrei Maltsev, *University of Maryland*
Abraham Mantell, *Nassau Community College*
Madhu Nayakkankuppam, *University of Maryland-Baltimore County*
Lei Ni, *Stanford University*
Gleb Novitchkov, *Penn State University*
Ralph Oberste-Vorth, *University of South Florida*
Dev Sinha, *Brown University*
Wasin So, *San Jose State University*
Ron Solomon, *Ohio State University*
Eugene Spiegel, *University of Connecticut*
Alan Stein, *University of Connecticut*
James Thomas, *Colorado State University*
Brian Turnquist, *Bethel College*
Michael Ward, *Western Oregon University*
Bruno Welfert, *Arizona State University*
Jack Xin, *University of Texas at Austin*

Para esta actualización de la tercera edición, agradezco a Thomas Polaski, de *Winthrop University*, quien revisó materiales complementarios de la tercera edición y siempre estuvo dispuesto a dar un consejo. También estoy agradecido con Rick Smith, de *University of Florida*, por adaptar sus proyectos de MATLAB para la actualización, y con Jeremy Case, de *Taylor University*, por su ayuda con los proyectos. Por último, agradezco a todo el personal de Addison-Wesley por su trabajo en esta actualización.

David C. Lay

Nota para los estudiantes

Este curso puede ser el más interesante y valioso entre todas las clases de matemáticas que pueden cursarse durante los estudios universitarios. De hecho, algunos estudiantes me han escrito o hablado después de graduarse y aún utilizan de manera ocasional este texto como una referencia en sus carreras en varias corporaciones importantes y en escuelas de posgrado en ingeniería. Los siguientes comentarios ofrecen algunos consejos prácticos e información que pueden ayudarle a dominar el material y a disfrutar el curso.

En álgebra lineal, los *conceptos* son tan importantes como los cálculos. Los ejercicios numéricos simples que inician cada serie de ejercicios sólo ayudan a verificar su comprensión de los procedimientos básicos. Posteriormente, en su carrera, las computadoras realizarán los cálculos, pero será necesario elegir los adecuados, saber cómo interpretar los resultados, y después explicar las soluciones a otras personas. Por esta razón, en el texto muchos ejercicios le piden explicar o justificar los cálculos realizados. Con frecuencia se solicita una explicación escrita como parte de la respuesta. Para la gran mayoría de los ejercicios con número impar, encontrará la explicación deseada o al menos una buena sugerencia. Debe evitar la tentación de buscar las respuestas a los ejercicios hasta no haber intentado escribir una solución por usted mismo. De otra manera, es posible considerar que algo ha sido comprendido aún cuando en realidad no sea así.

Para dominar los conceptos del álgebra lineal, es necesario leer y releer el texto con sumo cuidado. Los términos nuevos se presentan en negritas, algunas veces encerrados en recuadros de definición. Al final del texto se incluye un glosario de términos. Los conceptos importantes se establecen como teoremas o se incluyen en recuadros iluminados, para utilizarse como referencia rápida. Es recomendable leer las cuatro primeras páginas del prefacio para aprender más sobre la estructura del texto. Esto le proporcionará un marco para comprender la manera en que se desarrollará el curso.

En sentido práctico, el álgebra lineal es un lenguaje. Este lenguaje debe aprenderse de la misma forma en que se aprende un idioma extranjero —con trabajo diario—. El material presentado en una sección no se comprende con facilidad a menos que se haya estudiado por completo el texto y se hayan resuelto los ejercicios de las secciones previas. Por eso es necesario mantenerse al corriente con el curso, lo cual le ahorrará mucho tiempo y angustia.

Notas numéricas

Se recomienda leer las notas numéricas incluidas en el texto, incluso si no se está utilizando una computadora o calculadora gráfica junto con el libro. En la vida real, la mayor parte de las aplicaciones de álgebra lineal implican cálculos que están sujetos a algún error numérico, aún cuando dicho error pueda ser muy pequeño. Las notas numéricas le advertirán acerca de dificultades potenciales al utilizar posteriormente el álgebra lineal en su carrera, y si estudia estas notas ahora, existe una mayor posibilidad de que las recuerde después.

Si el lector disfruta la lectura de las notas numéricas, es posible que luego desee tomar un curso de álgebra numérica. Debido a la alta demanda de mayor poder computacional, los científicos en computación y los matemáticos trabajan en el álgebra lineal numérica para desarrollar algoritmos más rápidos y confiables con qué realizar cálculos, y los ingenieros eléctricos diseñan computadoras más rápidas y pequeñas para ejecutar los algoritmos. Este campo resulta estimulante, y su primer curso en álgebra lineal lo ayudará a prepararse para abordarlo.

Álgebra lineal numérica

Ciencias físicas

Estadística

1

Ecuaciones lineales en álgebra lineal

EJEMPLO INTRODUCTORIO

Modelos lineales en economía e ingeniería

A finales del verano de 1949 Wassily Leontief, profesor de Harvard, introdujo cuidadosamente la última de sus tarjetas perforadas en la computadora de la universidad, la Mark II. Las tarjetas contenían información acerca de la economía de Estados Unidos, y representaban un resumen de más de 250,000 piezas de información producidas por la oficina encargada de las estadísticas laborales en Estados Unidos después de dos años de trabajo intenso. Leontief había dividido la economía de Estados Unidos en 500 "sectores", tales como la industria del carbón, la industria automotriz, las comunicaciones, etc. Para cada sector, escribió una ecuación lineal que describía la forma en que dicho sector distribuía sus salidas hacia otros sectores de la economía. Debido a que la Mark II, una de las computadoras más grandes de la época, no podía manejar el sistema resultante de 500 ecuaciones y 500 incógnitas, Leontief había condensado el problema en un sistema de 42 ecuaciones y 42 incógnitas.

La programación de la computadora Mark II para las 42 ecuaciones de Leontief requirió varios meses de esfuerzo, y él estaba ansioso por ver cuánto tiempo le tomaría a la máquina resolver el problema. La Mark II zumbó y destelló durante 56 horas hasta que finalmente produjo una solución. La naturaleza de esta solución se analizará en las secciones 1.6 y 2.6.

Leontief, quien recibió el Premio Nobel de Economía en 1973, abrió la puerta a una nueva era en el modelado matemático de la economía. Sus esfuerzos desplegados en Harvard en 1949 marcaron uno de los primeros usos significativos de las computadoras para analizar lo que entonces era un modelo matemático a gran escala. Desde entonces, los investigadores de muchos otros campos han empleado computadoras para analizar modelos matemáticos. Debido a las masivas cantidades de datos involucrados, por lo general, los modelos son *lineales*; esto es, se describen mediante *sistemas de ecuaciones lineales*.

La importancia del álgebra lineal para las aplicaciones se ha elevado en proporción directa al aumento del poder de las computadoras, cada nueva generación de equipo y programas de cómputo dispara una demanda de capacidades aún mayores.

Por lo tanto, la ciencia de las computadoras está sólidamente ligada al álgebra lineal mediante el crecimiento explosivo de los procesamientos paralelos de datos y los cálculos a gran escala.

Los científicos e ingenieros trabajan ahora en problemas mucho más complejos de lo que creían posible hace unas cuantas décadas. En la actualidad, el álgebra lineal tiene para los estudiantes universitarios un mayor valor potencial en muchos campos científicos y de negocios que cualquier otra materia de matemáticas. El material incluido en este texto proporciona la base para un trabajo posterior en muchas áreas interesantes. A continuación se presentan unas cuantas posibilidades; posteriormente se describirán otras.

- *Exploración petrolera*. Cuando un barco busca depósitos submarinos de petróleo, *diariamente* sus computadoras resuelven miles de sistemas de ecuaciones lineales por separado. La información sísmica para elaborar las ecuaciones se obtiene a partir de ondas de choque submarinas creadas mediante explosiones con pistolas de aire. Las ondas rebotan en las rocas que hay bajo la superficie marina y se miden empleando geófonos conectados a extensos cables instalados debajo del barco.
- *Programación lineal*. En la actualidad, muchas decisiones administrativas importantes se toman con base en modelos de programación lineal que utilizan cientos de variables. Por ejemplo, la industria de las aerolíneas emplea programas lineales para crear los itinerarios de las tripulaciones de vuelo, monitorear las ubicaciones de los aviones, o planear los diversos programas de servicios de apoyo como mantenimiento y operaciones en terminal.
- *Redes eléctricas*. Los ingenieros utilizan programas de cómputo de simulación para diseñar circuitos eléctricos y microchips que incluyen millones de transistores. Estos programas utilizan técnicas de álgebra lineal y sistemas de ecuaciones lineales.

Los sistemas de ecuaciones lineales se encuentran en el corazón del álgebra lineal, y este capítulo los utiliza para introducir algunos de los conceptos centrales del álgebra lineal de una manera simple y concreta. En las secciones 1.1 y 1.2 se presenta un método sistemático para resolver sistemas de ecuaciones lineales. Este algoritmo se utilizará para realizar cálculos a lo largo del texto. En las secciones 1.3 y 1.4 se muestra cómo un sistema de ecuaciones lineales es equivalente a una *ecuación vectorial* y a una *ecuación matricial*. Esta equivalencia reducirá problemas que involucran combinaciones lineales de vectores a preguntas sobre los sistemas de ecuaciones lineales. Los conceptos fundamentales de generación, independencia lineal y transformaciones lineales, que se estudian en la segunda mitad del capítulo, desempeñarán un papel esencial a lo largo del texto mientras se explora la belleza y el poder del álgebra lineal.

1.1 SISTEMAS DE ECUACIONES LINEALES

Una **ecuación lineal** en las variables $x_1, \ldots, x_n$ es una ecuación que puede escribirse de la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \qquad (1)$$

donde b y los **coeficientes** $a_1, \ldots, a_n$ son números reales o complejos, por lo general conocidos. El subíndice n puede ser cualquier entero positivo. En los ejemplos y ejercicios del libro, n está normalmente entre 2 y 5. En los problemas de la vida real, n puede ser igual a 50, 5000, o incluso a valores más grandes.

Las ecuaciones

$$4x_1 - 5x_2 + 2 = x_1 \quad \text{y} \quad x_2 = 2\left(\sqrt{6} - x_1\right) + x_3$$

son ambas lineales porque pueden reordenarse algebraicamente como en la ecuación (1):

$$3x_1 - 5x_2 = -2 \quad \text{y} \quad 2x_1 + x_2 - x_3 = 2\sqrt{6}$$

Las ecuaciones

$$4x_1 - 5x_2 = x_1x_2 \quad \text{y} \quad x_2 = 2\sqrt{x_1} - 6$$

no son lineales debido a la presencia de x_1x_2 en la primera ecuación y $\sqrt{x_1}$ en la segunda.

Un **sistema de ecuaciones lineales** (o **sistema lineal**) es una colección de una o más ecuaciones lineales que involucran las mismas variables —digamos, $x_1, \ldots, x_n$. Un ejemplo es

$$\begin{aligned} 2x_1 - x_2 + 1.5x_3 &= 8 \\ x_1 \qquad\quad - 4x_3 &= -7 \end{aligned} \tag{2}$$

Una **solución** del sistema es una lista $(s_1, s_2, \ldots, s_n)$ de números que hacen de cada ecuación un enunciado verdadero cuando los valores $s_1, \ldots, s_n$ sustituyen, respectivamente, a $x_1, \ldots, x_n$. Por ejemplo, (5, 6.5, 3) es una solución del sistema (2) porque, cuando estos valores sustituyen en (2) a x_1, x_2 y x_3, respectivamente, las ecuaciones se simplifican a $8 = 8$ y $-7 = -7$.

El conjunto de todas las soluciones posibles se llama **conjunto solución** del sistema lineal. Se dice que dos sistemas lineales son **equivalentes** si tienen el mismo conjunto solución. Esto es, cada solución del primer sistema es una solución del segundo sistema, y cada solución del segundo sistema es una solución del primero.

Determinar el conjunto solución de un sistema de dos ecuaciones lineales resulta sencillo porque consiste en localizar la intersección de dos rectas. Un problema típico es

$$\begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned}$$

Las gráficas de estas ecuaciones son rectas, las cuales se denotan mediante ℓ_1 y ℓ_2. Un par de números (x_1, x_2) satisface las *dos* ecuaciones de este sistema si, y sólo si, el punto (x_1, x_2) pertenece tanto a ℓ_1 como a ℓ_2. En el sistema anterior, la solución es el punto único (3, 2), lo cual puede verificarse con facilidad. Vea la figura 1.

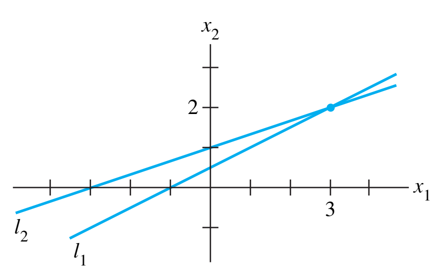


FIGURA 1 Exactamente una solución.

Por supuesto, la intersección de dos rectas no debe darse necesariamente en un solo punto —las rectas pueden ser paralelas o coincidir y, por lo tanto, “intersecar” en todos los puntos sobre la recta. En la figura 2 se muestran las gráficas que corresponden a los siguientes sistemas:

$$\text{(a)}\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned} \qquad \text{(b)}\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 1 \end{aligned}$$

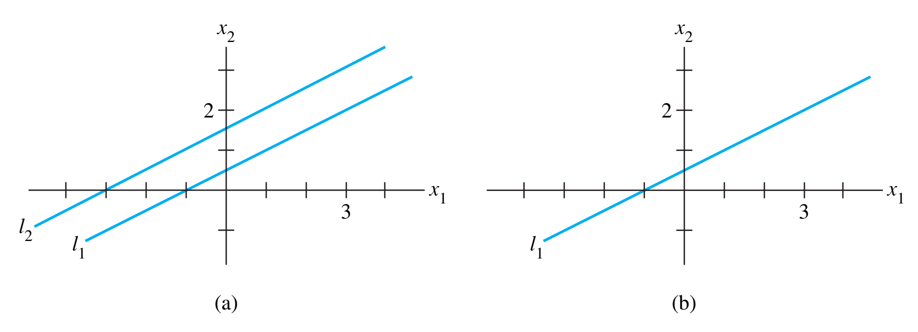


FIGURA 2 (a) Sin solución. (b) Con infinidad de soluciones.

Las figuras 1 y 2 ilustran los siguientes hechos generales acerca de los sistemas lineales, los cuales serán verificados en la sección 1.2.

Un sistema de ecuaciones lineales puede

1. no tener solución, o
2. tener exactamente una solución, o
3. tener una cantidad infinita de soluciones.

Se dice que un sistema de ecuaciones lineales es **consistente** si tiene una solución o una infinidad de soluciones; un sistema es **inconsistente** cuando no tiene ninguna solución.

Notación matricial

La información esencial de un sistema lineal puede registrarse de manera compacta en un arreglo rectangular llamado **matriz**. Dado el sistema

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad (3)$$

con los coeficientes de cada variable alineados en columnas, la matriz

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

se denomina **matriz coeficiente** (o **matriz de coeficientes**) del sistema (3), y

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \tag{4}$$

se denomina **matriz aumentada** del sistema. (Aquí, la segunda fila contiene un cero porque la segunda ecuación podría escribirse como $0 \cdot x_1 + 2x_2 - 8x_3 = 8$.) La matriz aumentada de un sistema consta de su matriz de coeficientes con una columna adicional que contiene las constantes de los lados derechos de las ecuaciones.

El **tamaño** de una matriz indica el número de filas y columnas que la integran. La matriz aumentada (4) que se presentó líneas arriba tiene 3 filas y 4 columnas y se conoce como una matriz de 3×4 (se lee "3 por 4"). Si m y n son enteros positivos, una **matriz** $\boldsymbol{m \times n}$ es un arreglo rectangular de números con m filas y n columnas. (El número de filas siempre va primero.) La notación matricial simplificará los cálculos de los ejemplos que se presentan enseguida.

Resolución de un sistema lineal

En esta sección y en la siguiente se describe un algoritmo, o procedimiento sistemático, para resolver sistemas lineales. La estrategia básica es *reemplazar un sistema con un sistema equivalente (es decir, uno con el mismo conjunto solución) que sea más fácil de resolver.*

Dicho de manera sencilla, utilice el término x_1 que esté presente en la primera ecuación de un sistema para eliminar los términos x_1 que haya en las otras ecuaciones. Después use el término x_2 presente en la segunda ecuación para eliminar los términos x_2 en las otras ecuaciones, y así sucesivamente, hasta que obtenga un sistema de ecuaciones equivalente muy simple.

Para simplificar un sistema lineal se utilizan tres operaciones básicas: reemplazar una ecuación mediante la suma de la propia ecuación y un múltiplo de otra ecuación, intercambiar dos ecuaciones, y multiplicar todos los términos de una ecuación por una constante distinta de cero. Después del primer ejemplo, se verá por qué estas tres operaciones no cambian el conjunto solución del sistema.

EJEMPLO 1 Resuelva el sistema (3).

Solución El procedimiento de eliminación se muestra enseguida con y sin notación matricial, y los resultados se colocan uno junto al otro para compararlos:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

Mantenga x_1 en la primera ecuación y elimínela de las otras ecuaciones. Para hacer esto, sume 4 veces la ecuación 1 a la ecuación 3. Por lo general, luego de alguna práctica este tipo de cálculos se realizan mentalmente:

$$\begin{array}{rr} 4 \cdot [\text{ecuación 1}]: & 4x_1 - 8x_2 + 4x_3 = 0 \\ + [\text{ecuación 3}]: & -4x_1 + 5x_2 + 9x_3 = -9 \\ \hline [\text{nueva ecuación 3}]: & -3x_2 + 13x_3 = -9 \end{array}$$

El resultado de este cálculo se escribe en lugar de la tercera ecuación original:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

Ahora, multiplique la ecuación 2 por $1/2$ para obtener 1 como el coeficiente para x_2. (Este cálculo simplificará la aritmética del siguiente paso.)

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

Utilice x_2 en la ecuación 2 para eliminar $-3x_2$ en la ecuación 3. El cálculo "mental" es

$$\begin{array}{rr} 3\cdot[\text{ecuación } 2]: & 3x_2 - 12x_3 = 12 \\ +\,[\text{ecuación } 3]: & -3x_2 + 13x_3 = -9 \\ \hline [\text{nueva ecuación } 3]: & x_3 = 3 \end{array}$$

El nuevo sistema tiene una forma *triangular*:[1]

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Al final, se deseará eliminar el término $-2x_2$ de la ecuación 1, pero resulta más eficiente utilizar primero x_3 en la ecuación 3, para eliminar los términos $-4x_3$ y $+x_3$ en las ecuaciones 2 y 1. Los dos cálculos "mentales" son

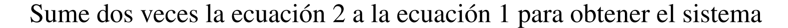
$$\begin{array}{rr} 4\cdot[\text{ec. } 3]: & 4x_3 = 12 \\ +\,[\text{ec. } 2]: & x_2 - 4x_3 = 4 \\ \hline [\text{nueva ec. } 2]: & x_2 \qquad = 16 \end{array} \qquad \begin{array}{rr} -1\cdot[\text{ec. } 3]: & -x_3 = -3 \\ +\,[\text{ec. } 1]: & x_1 - 2x_2 + x_3 = 0 \\ \hline [\text{nueva ec. } 1]: & x_1 - 2x_2 \qquad = -3 \end{array}$$

Es conveniente combinar los resultados de estas dos operaciones:

$$\begin{aligned} x_1 - 2x_2 \qquad &= -3 \\ x_2 \qquad &= 16 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

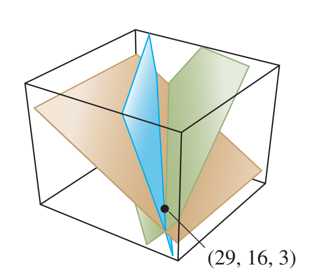
Ahora, después de haber limpiado la columna que está sobre la x_3 en la ecuación 3, regrese a la x_2 en la ecuación 2 y úsela para eliminar el $-2x_2$ ubicado sobre ella. Debido al trabajo previo realizado con x_3, ahora no existe ninguna operación que involucre a términos de x_3.

[1]En la próxima sección, el término intuitivo *triangular* se reemplazará por uno más preciso.



Sume dos veces la ecuación 2 a la ecuación 1 para obtener el sistema

$$\begin{cases} x_1 \qquad\qquad = 29 \\ \qquad x_2 \qquad = 16 \\ \qquad\qquad x_3 = 3 \end{cases} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$



Cada una de las ecuaciones originales determina un plano en el espacio tridimensional. El punto (29, 16, 3) pertenece a los tres planos.

En esencia, el trabajo ya está hecho. Se observa que la solución única del sistema original es (29, 16, 3). Sin embargo, como hay muchos cálculos involucrados, resulta una buena práctica verificar las operaciones. Para comprobar que (29, 16, 3) *es* una solución, sustituya estos valores en el lado izquierdo del sistema original, y calcule:

$$\begin{aligned} (29) - 2(16) + (3) &= 29 - 32 + 3 = 0 \\ 2(16) - 8(3) &= 32 - 24 = 8 \\ -4(29) + 5(16) + 9(3) &= -116 + 80 + 27 = -9 \end{aligned}$$

Los resultados coinciden con el lado derecho del sistema original, así que (29, 16, 3) es una solución del sistema.

En el ejemplo 1 se ilustra cómo, en un sistema lineal, las operaciones sobre ecuaciones corresponden a las operaciones sobre las filas apropiadas de la matriz aumentada. Las tres operaciones básicas mencionadas con anterioridad corresponden a las siguientes operaciones sobre la matriz aumentada.

OPERACIONES ELEMENTALES DE FILA

1. (Reemplazo) Reemplazar una fila por la suma de sí misma y un múltiplo de otra fila.[2]
2. (Intercambio) Intercambiar dos filas.
3. (Escalamiento) Multiplicar todas las entradas de una fila por una constante distinta de cero.

Las operaciones de fila pueden aplicarse a cualquier matriz, no únicamente a una que surja como la matriz aumentada de un sistema lineal. Se dice que dos matrices son **equivalentes por filas** si existe una sucesión de operaciones elementales de fila que convierta una matriz en la otra.

Es importante advertir que las operaciones de fila son *reversibles*. Si dos filas se intercambian, pueden regresarse a sus posiciones originales mediante otro intercambio. Si una fila se escala mediante una constante c distinta de cero, al multiplicar después la nueva fila por $1/c$ se obtiene la fila original. Por último, considere una operación de reemplazo que involucra dos filas —por ejemplo, las filas 1 y 2— y suponga que a la fila 2 se le suma la fila 1 multiplicada por c para producir un nueva fila 2. Si desea "revertir" esta operación, sume a la nueva fila 2 la fila 1 multiplicada por $-c$ y obtenga la fila 2 original. Vea los ejercicios 29 a 32 al final de esta sección.

Por el momento, nuestro interés reside en las operaciones de fila sobre la matriz aumentada de un sistema de ecuaciones lineales. Suponga un sistema que se transforma en otro nuevo mediante operaciones de fila.

[2]Una paráfrasis común del reemplazo de una fila es "sumar a una fila un múltiplo de otra fila".

Al considerar cada uno de los tipos de operaciones de fila, puede advertirse que cualquier solución del sistema original continúa siendo una solución del sistema nuevo. Asimismo, como el sistema original puede producirse mediante operaciones de fila sobre el sistema nuevo, cada una de las soluciones del sistema nuevo también es una solución del sistema original. Esta explicación justifica el hecho siguiente.

Si las matrices aumentadas de dos sistemas lineales son equivalentes por filas, entonces los dos sistemas tienen el mismo conjunto solución.

Aunque el ejemplo 1 es extenso, puede afirmarse que, después de algún tiempo de práctica, los cálculos se ejecutan con rapidez. Por lo general, en el texto y en los ejercicios las operaciones de fila serán muy fáciles de realizar, lo cual permitirá que el estudiante se enfoque en los conceptos importantes. No obstante, se recomienda aprender a realizar operaciones de fila de manera precisa porque se utilizarán a lo largo de todo el libro.

En el resto de esta sección se muestra cómo utilizar las operaciones de fila para determinar el tamaño de un conjunto solución, sin resolver por completo el sistema lineal.

Preguntas de existencia y unicidad

En la sección 1.2 se estudiará porqué un conjunto solución para un sistema lineal puede no contener ninguna solución, contener solamente una solución, o contener una infinidad de soluciones. Para determinar cuál posibilidad es verdadera para un sistema en particular, se formulan dos preguntas.

DOS PREGUNTAS FUNDAMENTALES ACERCA DE UN SISTEMA LINEAL

1. ¿El sistema es consistente? Es decir, ¿*existe* al menos una solución?
2. Si existe solución, ¿*sólo* hay una? Esto es, ¿la solución es *única*?

Estas dos preguntas aparecerán a lo largo del texto en muchas formas diferentes. En esta sección y en la próxima, se mostrará cómo contestarlas mediante operaciones de fila sobre la matriz aumentada.

EJEMPLO 2 Determine si el siguiente sistema es consistente:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

Solución Éste es el sistema del ejemplo 1. Suponga que se realizan las operaciones necesarias para obtener la forma triangular

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

En este punto ya se conoce x_3; si su valor se sustituyera en la ecuación 2, sería posible calcular x_2 y, por ende, se podría determinar x_1 a partir de la ecuación 1. Por lo tanto,

existe una solución; y el sistema es consistente. (De hecho, x_2 se determina únicamente con la ecuación 2 puesto que x_3 tiene un solo valor posible, y por lo tanto x_1 se resuelve solamente a partir de la ecuación 1. De manera que la solución es única.)

EJEMPLO 3 Determine si el siguiente sistema es consistente:

$$\begin{aligned} x_2 - 4x_3 &= 8 \\ 2x_1 - 3x_2 + 2x_3 &= 1 \\ 5x_1 - 8x_2 + 7x_3 &= 1 \end{aligned} \tag{5}$$

Solución La matriz aumentada es

$$\begin{bmatrix} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

Para obtener una x_1 en la primera ecuación, se intercambian las filas 1 y 2:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

Para eliminar el término $5x_1$ en la tercera ecuación, se agrega a la fila 3 la fila 1 multiplicada por $-5/2$:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & -1/2 & 2 & -3/2 \end{bmatrix} \tag{6}$$

Enseguida, utilice el término x_2 en la segunda ecuación para eliminar el término $-(1/2)x_2$ de la tercera ecuación. Sume a la fila 3 la fila 2 multiplicada por $1/2$:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \tag{7}$$

Ahora, la matriz aumentada está en forma triangular. Para interpretarla de manera correcta, regrese a la notación con ecuaciones:

$$\begin{aligned} 2x_1 - 3x_2 + 2x_3 &= 1 \\ x_2 - 4x_3 &= 8 \\ 0 &= 5/2 \end{aligned} \tag{8}$$

Este sistema es inconsistente porque no existe un punto que pertenezca de manera simultánea a los tres planos.

La ecuación $0 = 5/2$ es una forma corta de $0x_1 + 0x_2 + 0x_3 = 5/2$. Desde luego, este sistema en forma triangular tiene una contradicción. No existen valores de x_1, x_2, x_3 que satisfagan (8) porque la ecuación $0 = 5/2$ nunca es verdadera. Como (8) y (5) tienen el mismo conjunto solución, el sistema original es inconsistente (es decir, no tiene solución).

Preste atención especial a la matriz aumentada en (7). Su última fila es típica de un sistema inconsistente en forma triangular.

NOTA NUMÉRICA

En problemas reales, los sistemas de ecuaciones lineales se resuelven empleando una computadora. Para una matriz de coeficientes cuadrada, los programas de cómputo casi siempre usan el algoritmo de eliminación que se presenta aquí en la sección 1.2, con pequeñas modificaciones para mejorar su precisión.

La gran mayoría de los problemas de álgebra lineal que se presentan en los negocios y la industria se resuelven con programas que utilizan la *aritmética de punto flotante*. Los números se representan como decimales $\pm .d_1 \cdots d_p \times 10^r$, donde r es un entero y el número p de dígitos a la derecha del punto decimal usualmente se encuentra entre 8 y 16. Normalmente, las operaciones aritméticas con estos números resultan inexactas, porque el resultado debe redondearse (o truncarse) al número de dígitos almacenados. El "error de redondeo" también se presenta cuando un número como 1/3 es introducido a la computadora, puesto que su representación debe aproximarse mediante un número finito de dígitos. Por fortuna, las inexactitudes de la aritmética de punto flotante muy pocas veces causan problemas. Las notas numéricas incluidas en este libro lo prevendrán, ocasionalmente, sobre aspectos que podrá necesitar tener en consideración más adelante en su carrera.

PROBLEMAS DE PRÁCTICA

A lo largo del texto, debe intentar resolver los problemas de práctica antes de trabajar con los ejercicios. Después de cada serie de ejercicios se presentan las soluciones.

1. Exprese con sus propias palabras la siguiente operación elemental de fila que debe realizarse para resolver los sistemas presentados a continuación. [Para (*a*), existe más de una respuesta posible.]

a.
$$\begin{aligned} x_1 + 4x_2 - 2x_3 + 8x_4 &= 12 \\ x_2 - 7x_3 + 2x_4 &= -4 \\ 5x_3 - x_4 &= 7 \\ x_3 + 3x_4 &= -5 \end{aligned}$$

b.
$$\begin{aligned} x_1 - 3x_2 + 5x_3 - 2x_4 &= 0 \\ x_2 + 8x_3 \qquad &= -4 \\ 2x_3 \qquad &= 3 \\ x_4 &= 1 \end{aligned}$$

2. La matriz aumentada de un sistema lineal ha sido transformada mediante operaciones de fila a la forma que se presenta a continuación. Determine si el sistema es consistente.

$$\begin{bmatrix} 1 & 5 & 2 & -6 \\ 0 & 4 & -7 & 2 \\ 0 & 0 & 5 & 0 \end{bmatrix}$$

3. ¿Es (3, 4, −2) una solución del siguiente sistema?

$$\begin{aligned} 5x_1 - x_2 + 2x_3 &= 7 \\ -2x_1 + 6x_2 + 9x_3 &= 0 \\ -7x_1 + 5x_2 - 3x_3 &= -7 \end{aligned}$$

4. ¿Para cuáles valores de h y k es consistente el siguiente sistema?

$$\begin{aligned} 2x_1 - x_2 &= h \\ -6x_1 + 3x_2 &= k \end{aligned}$$

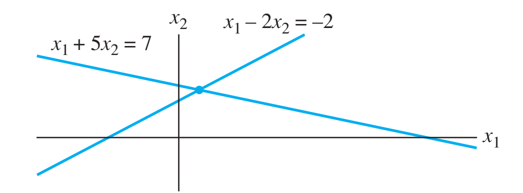
1.1 EJERCICIOS

Resuelva los sistemas de los ejercicios 1 a 4 usando las operaciones elementales de fila sobre las ecuaciones o sobre la matriz aumentada. Utilice el procedimiento de eliminación sistemática descrito en esta sección.

1. $$\begin{aligned} x_1 + 5x_2 &= 7 \\ -2x_1 - 7x_2 &= -5 \end{aligned}$$

2. $$\begin{aligned} 2x_1 + 4x_2 &= -4 \\ 5x_1 + 7x_2 &= 11 \end{aligned}$$

3. Encuentre el punto (x_1, x_2) que pertenece tanto a la línea $x_1 + 5x_2 = 7$ como a la línea $x_1 - 2x_2 = -2$. Vea la figura.



4. Encuentre el punto de intersección de las rectas $x_1 - 5x_2 = 1$ y $3x_1 - 7x_2 = 5$.

Considere cada matriz de los ejercicios 5 y 6 como la matriz aumentada de un sistema lineal. Exprese con sus propias palabras las siguientes dos operaciones elementales de fila que deben realizarse en el proceso para resolver el sistema.

5. $$\begin{bmatrix} 1 & -4 & 5 & 0 & 7 \\ 0 & 1 & -3 & 0 & 6 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -5 \end{bmatrix}$$

6. $$\begin{bmatrix} 1 & -6 & 4 & 0 & -1 \\ 0 & 2 & -7 & 0 & 4 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 3 & 1 & 6 \end{bmatrix}$$

En los ejercicios 7 a 10, la matriz aumentada de un sistema lineal ha sido reducida mediante operaciones de fila a la forma que se muestra. En cada caso, ejecute las operaciones de fila apropiadas y describa el conjunto solución del sistema original.

7. $$\begin{bmatrix} 1 & 7 & 3 & -4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

8. $$\begin{bmatrix} 1 & -4 & 9 & 0 \\ 0 & 1 & 7 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$

9. $$\begin{bmatrix} 1 & -1 & 0 & 0 & -4 \\ 0 & 1 & -3 & 0 & -7 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 2 & 4 \end{bmatrix}$$

10. $$\begin{bmatrix} 1 & -2 & 0 & 3 & -2 \\ 0 & 1 & 0 & -4 & 7 \\ 0 & 0 & 1 & 0 & 6 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix}$$

Resuelva los sistemas de los ejercicios 11 a 14.

11. $$\begin{aligned} x_2 + 4x_3 &= -5 \\ x_1 + 3x_2 + 5x_3 &= -2 \\ 3x_1 + 7x_2 + 7x_3 &= 6 \end{aligned}$$

12. $$\begin{aligned} x_1 - 3x_2 + 4x_3 &= -4 \\ 3x_1 - 7x_2 + 7x_3 &= -8 \\ -4x_1 + 6x_2 - x_3 &= 7 \end{aligned}$$

13. $$\begin{aligned} x_1 \qquad - 3x_3 &= 8 \\ 2x_1 + 2x_2 + 9x_3 &= 7 \\ x_2 + 5x_3 &= -2 \end{aligned}$$

14. $$\begin{aligned} x_1 - 3x_2 \qquad &= 5 \\ -x_1 + x_2 + 5x_3 &= 2 \\ x_2 + x_3 &= 0 \end{aligned}$$

Determine si los sistemas de los ejercicios 15 y 16 son consistentes. No resuelva los sistemas por completo.

15. $$\begin{aligned} x_1 \qquad + 3x_3 \qquad &= 2 \\ x_2 \qquad - 3x_4 &= 3 \\ -2x_2 + 3x_3 + 2x_4 &= 1 \\ 3x_1 \qquad\qquad + 7x_4 &= -5 \end{aligned}$$

16. $$\begin{aligned} x_1 \qquad\qquad - 2x_4 &= -3 \\ 2x_2 + 2x_3 \qquad &= 0 \\ x_3 + 3x_4 &= 1 \\ -2x_1 + 3x_2 + 2x_3 + x_4 &= 5 \end{aligned}$$

17. ¿Las tres rectas $x_1 - 4x_2 = 1$, $2x_1 - x_2 = -3$, y $-x_1 - 3x_2 = 4$ tienen un punto de intersección común? Explique su respuesta.

18. ¿Los tres planos $x_1 + 2x_2 + x_3 = 4$, $x_2 - x_3 = 1$, y $x_1 + 3x_2 = 0$ tienen al menos un punto de intersección común? Explique su respuesta.

En los ejercicios 19 a 22, determine el valor o los valores de h tales que la matriz dada es la matriz aumentada de un sistema lineal consistente.

19. $\begin{bmatrix} 1 & h & 4 \\ 3 & 6 & 8 \end{bmatrix}$

20. $\begin{bmatrix} 1 & h & -3 \\ -2 & 4 & 6 \end{bmatrix}$

21. $\begin{bmatrix} 1 & 3 & -2 \\ -4 & h & 8 \end{bmatrix}$

22. $\begin{bmatrix} 2 & -3 & h \\ -6 & 9 & 5 \end{bmatrix}$

En los ejercicios 23 y 24, varios enunciados clave de esta sección se citan directamente, se han modificado un poco (pero siguen siendo verdaderos), o se han alterado de alguna forma que los vuelve falsos en algunos casos. Marque cada enunciado como verdadero o falso y *justifique* su respuesta. (Si el enunciado es verdadero, dé la ubicación aproximada en el texto donde aparece uno similar o haga referencia a una definición o teorema. Si es falso, dé la ubicación del enunciado que se cita o utiliza de manera incorrecta, o proporcione un ejemplo que muestre que no es verdadero en todos los casos.) En muchas secciones de este texto aparecerán preguntas similares del tipo verdadero/falso.

23. a. Todas las operaciones elementales de fila son reversibles.

b. Una matriz de 5×6 tiene seis filas.

c. El conjunto solución de un sistema lineal que incluya las variables $x_1, \ldots, x_n$ es una lista de números $(s_1, \ldots, s_n)$ que hace de cada ecuación del sistema un enunciado verdadero cuando los valores $s_1, \ldots, s_n$ sustituyen, respectivamente, a $x_1, \ldots, x_n$.

d. Las dos preguntas fundamentales acerca de un sistema lineal involucran la existencia y la unicidad.

24. a. En una matriz aumentada, las operaciones elementales de fila no cambian nunca el conjunto solución del sistema lineal asociado.

b. Dos matrices son equivalentes por filas cuando poseen el mismo número de filas.

c. Un sistema inconsistente tiene más de una solución.

d. Dos sistemas lineales son equivalentes si tienen el mismo conjunto solución.

25. Encuentre una ecuación que involucre a g, h y k, la cual permita que esta matriz aumentada corresponda a un sistema consistente:

$$\begin{bmatrix} 1 & -4 & 7 & g \\ 0 & 3 & -5 & h \\ -2 & 5 & -9 & k \end{bmatrix}$$

26. Construya tres matrices aumentadas diferentes de tres sistemas lineales cuyo conjunto solución sea $x_1 = -2, x_2 = 1, x_3 = 0$.

27. Suponga que el sistema presentado a continuación es consistente para todos los valores posibles de f y g. ¿Qué puede afirmarse acerca de los coeficientes c y d? Justifique su respuesta.

$$\begin{aligned} x_1 + 3x_2 &= f \\ cx_1 + dx_2 &= g \end{aligned}$$

28. Suponga que a, b, c y d son constantes de tal forma que a es diferente de cero y el sistema presentado a continuación es consistente para todos los valores posibles de f y g. ¿Qué puede afirmarse acerca de los números a, b, c y d? Justifique su respuesta.

$$\begin{aligned} ax_1 + bx_2 &= f \\ cx_1 + dx_2 &= g \end{aligned}$$

En los ejercicios 29 a 32, encuentre la operación elemental de fila que transforma la primera matriz en la segunda, determine entonces la operación de fila inversa que transforma la segunda matriz en la primera.

29. $\begin{bmatrix} 0 & -2 & 5 \\ 1 & 4 & -7 \\ 3 & -1 & 6 \end{bmatrix}, \begin{bmatrix} 1 & 4 & -7 \\ 0 & -2 & 5 \\ 3 & -1 & 6 \end{bmatrix}$

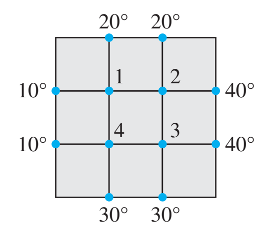
30. $\begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & -5 & 9 \end{bmatrix}, \begin{bmatrix} 1 & 3 & -4 \\ 0 & 1 & -3 \\ 0 & -5 & 9 \end{bmatrix}$

31. $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -2 & 8 \\ 4 & -1 & 3 & -6 \end{bmatrix}, \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -2 & 8 \\ 0 & 7 & -1 & -6 \end{bmatrix}$

32. $\begin{bmatrix} 1 & 2 & -5 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & -3 & 9 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 2 & -5 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

Un aspecto importante en el estudio de la transferencia de calor es determinar la distribución de la temperatura en estado estable sobre una placa delgada cuando se conoce la temperatura presente alrededor de los bordes. Suponga que la placa mostrada en la figura representa la sección transversal de una viga de metal, con un flujo de calor insignificante en la dirección perpendicular a la placa. Sean $T_1, \ldots, T_4$ las temperaturas en los cuatro nodos interiores de la malla que se muestra en la figura. En un nodo, la temperatura es aproximadamente igual al promedio de los cuatro nodos más cercanos —a la izquierda, arriba, a la derecha y abajo.[3] Por ejemplo,

$$T_1 = (10 + 20 + T_2 + T_4)/4, \quad \text{o} \quad 4T_1 - T_2 - T_4 = 30$$



[3]Vea Frank M. White, *Heat and Mass Transfer* (Reading, MA: Addison-Wesley Publishing, 1991), pp. 145−149.

33. Escriba un sistema de cuatro ecuaciones cuya solución proporcione un estimado para las temperaturas $T_1, \ldots, T_4$.

34. Resuelva el sistema de ecuaciones del ejercicio 33. [*Sugerencia:* Para acelerar los cálculos, intercambie las filas 1 y 4 antes de comenzar las operaciones de "reemplazo".]

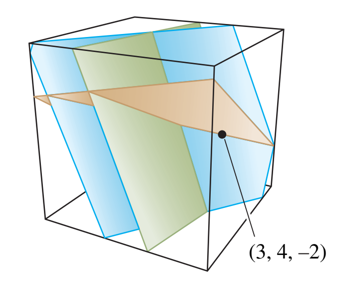
SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. a. Para realizar "cálculos a mano", lo mejor es intercambiar las ecuaciones 3 y 4. Otra posibilidad es multiplicar la ecuación 3 por $1/5$; o reemplazar la ecuación 4 por su suma con la fila 3 multiplicada por $-1/5$. (En cualquier caso, no utilice x_2 en la ecuación 2 para eliminar $4x_2$ en la ecuación 1. Espere hasta alcanzar la forma triangular y hasta que los términos con x_3 y x_4 hayan sido eliminados de las primeras dos ecuaciones.)

b. El sistema está en forma triangular. La simplificación posterior comienza con x_4 en la cuarta ecuación. Utilice esta x_4 para eliminar todos los términos con x_4 localizados arriba de ella. Ahora, el paso adecuado es sumar la ecuación 4, multiplicada por 2, con la ecuación 1. (Después de esto, vaya a la ecuación 3, multiplíquela por $1/2$, y utilice la ecuación resultante para eliminar los términos con x_3 ubicados arriba de ella.)

2. El sistema correspondiente a la matriz aumentada es

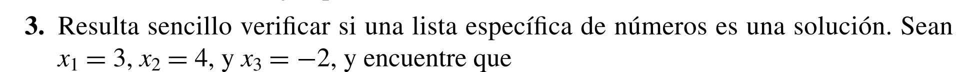
$$\begin{aligned} x_1 + 5x_2 + 2x_3 &= -6 \\ 4x_2 - 7x_3 &= 2 \\ 5x_3 &= 0 \end{aligned}$$



Como $(3, 4, -2)$ satisface las dos primeras ecuaciones, se encuentra sobre la línea de intersección de los dos primeros planos. Como $(3, 4, -2)$ no satisface las tres ecuaciones, no pertenece a los tres planos.

La tercera ecuación vuelve $x_3 = 0$, que ciertamente es un valor permisible para x_3. Después, al eliminar los términos con x_3 en las ecuaciones 1 y 2, es posible encontrar valores únicos para x_2 y x_1. Por lo tanto, existe una solución y es única. Compare esta situación con la del ejemplo 3.

3. Resulta sencillo verificar si una lista específica de números es una solución. Sean $x_1 = 3$, $x_2 = 4$, y $x_3 = -2$, y encuentre que



$$\begin{aligned} 5(3) - (4) + 2(-2) &= 15 - 4 - 4 = 7 \\ -2(3) + 6(4) + 9(-2) &= -6 + 24 - 18 = 0 \\ -7(3) + 5(4) - 3(-2) &= -21 + 20 + 6 = 5 \end{aligned}$$

Aunque se satisfacen las primeras dos ecuaciones, la tercera no, entonces $(3, 4, -2)$ no es una solución al sistema. Observe el uso de paréntesis cuando se hacen sustituciones, los cuales son muy recomendables como protección contra errores aritméticos.

4. Cuando la segunda ecuación se reemplaza por su suma con la primera ecuación multiplicada por 3, el sistema se convierte en

$$\begin{aligned} 2x_1 - x_2 &= h \\ 0 &= k + 3h \end{aligned}$$

Si $k + 3h$ es diferente de cero, el sistema no tiene solución. El sistema es consistente para cualesquiera valores de h y k que produzcan $k + 3h = 0$.

1.2 REDUCCIÓN POR FILAS Y FORMAS ESCALONADAS

En esta sección se perfecciona el método de la sección 1.1 en un algoritmo de reducción por filas que permitirá analizar cualquier sistema de ecuaciones lineales.[1] Las preguntas fundamentales de existencia y unicidad, expuestas en la sección 1.1, podrán contestarse utilizando la primera parte del algoritmo.

El algoritmo se aplica a cualquier matriz, ya sea vista como una matriz aumentada para un sistema lineal o no. Entonces, la primera parte de esta sección trata acerca de una matriz rectangular arbitraria. Se comienza por introducir dos clases importantes de matrices que incluyen las matrices "triangulares" de la sección 1.1. En las definiciones presentadas a continuación, una fila o una columna *distinta de cero* en una matriz serán una fila o una columna que contengan al menos una entrada diferente de cero; una **entrada principal** de una fila se refiere a la entrada diferente de cero que se encuentra más a la izquierda (en una fila distinta de cero).

DEFINICIÓN

Una matriz rectangular está en **forma escalonada** (o en **forma escalonada por filas**) si tiene las tres propiedades siguientes:

1. Todas las filas distintas de cero están arriba de cualquier fila integrada sólo por ceros.
2. Cada entrada principal de una fila está en una columna situada a la derecha de la entrada principal de la fila que se encuentra arriba de dicha entrada.
3. Todas las entradas que se localicen en una columna situada debajo de una entrada principal son ceros.

Si una matriz en forma escalonada satisface las siguientes condiciones adicionales, entonces se encuentra en **forma escalonada reducida** (o **forma escalonada reducida por filas**):

4. La entrada principal de cada fila distinta de cero es 1.
5. Cada 1 principal es la única entrada distinta de cero en su columna.

Una **matriz escalonada** (respectivamente, **matriz escalonada reducida**) es una matriz que está en forma escalonada (respectivamente, forma escalonada reducida). La propiedad 2 enuncia que las entradas principales forman un patrón *escalonado* ("como escalera") que avanza hacia abajo y a la derecha de la matriz. La propiedad 3 es una simple consecuencia de la propiedad 2, pero se incluyó aquí para enfatizarla.

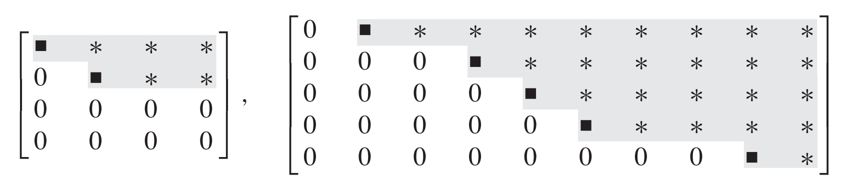
Las matrices "triangulares" de la sección 1.1, tales como

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

[1]Este algoritmo es una variación de lo que se conoce comúnmente como *eliminación gaussiana*. Los matemáticos chinos utilizaron un método de eliminación similar alrededor del año 250 a.C. El proceso no se conoció en la cultura occidental sino hasta el siglo XIX, cuando un famoso matemático alemán, Carl Friedrich Gauss, lo descubrió. Un ingeniero alemán, Wilhelm Jordan, popularizó el algoritmo al emplearlo en un texto sobre geodesia en 1888.

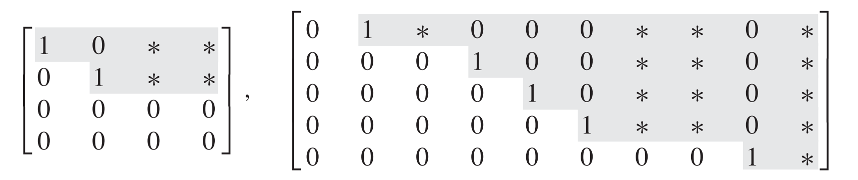
están en forma escalonada. De hecho, la segunda matriz está en forma escalonada reducida. A continuación se presentan ejemplos adicionales.

EJEMPLO 1 Las siguientes matrices están en forma escalonada. Las entradas principales (■) pueden tener cualquier valor distinto de cero; las entradas con asterisco (*) pueden tener cualquier valor (incluso cero).



$$\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \end{bmatrix}$$

Las siguientes matrices están en forma escalonada reducida porque las entradas principales son números 1, y abajo y arriba de cada 1 principal sólo existen ceros.



$$\begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

Cualquier matriz distinta de cero se puede **reducir por filas** (esto es, transformarse mediante operaciones elementales de fila) para producir más de una matriz en forma escalonada, para ello se usan diferentes sucesiones de operaciones de fila. Sin embargo, la forma escalonada reducida que se obtiene a partir de una matriz es única. El teorema siguiente se comprueba en el apéndice A incluido al final del texto.

TEOREMA 1 **Unicidad de la forma escalonada reducida**

Cada matriz es equivalente por filas a una y sólo una matriz escalonada reducida.

Si una matriz *A* es equivalente por filas a una matriz escalonada *U*, se dice que *U* es una **forma escalonada** (o una forma escalonada por filas) **de A**; si *U* está en su forma escalonada reducida, se afirma que es la **forma escalonada reducida de A**. [La mayoría de los programas de matrices y de las calculadoras con capacidad para resolver matrices utilizan la abreviatura RREF para encontrar la forma escalonada reducida (por filas). Algunos usan REF para la forma escalonada (por filas) (del inglés *row reduced echelon form* y *row echelon form*).]

Posiciones pivote

Cuando las operaciones de fila sobre una matriz producen una forma escalonada, las operaciones de fila posteriores para obtener la forma escalonada reducida no cambian las posiciones de las entradas principales. Como la forma escalonada reducida es única, *las entradas principales siempre están en las mismas posiciones en cualquier forma escalonada obtenida a partir de una matriz dada.* Estas entradas principales corresponden a los números 1 principales que hay en la forma escalonada reducida.

DEFINICIÓN En una matriz A, una **posición pivote** es una ubicación en A que corresponde a un 1 principal en la forma escalonada reducida de A. Una **columna pivote** es una columna de A que contiene una posición pivote.

En el ejemplo 1, los cuadros (■) identifican las posiciones pivote. Muchos conceptos fundamentales incluidos en los primeros cuatro capítulos de este libro estarán conectados de una forma u otra con las posiciones pivote que aparecen en una matriz.

EJEMPLO 2 Reduzca por filas la matriz A que se muestra a continuación hasta la forma escalonada, y localice las columnas pivote de A.

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

Solución Use la misma estrategia básica aplicada en la sección 1.1. El elemento superior de la columna distinta de cero que se encuentra más a la izquierda de la matriz es la primera posición pivote. En esta posición, debe colocarse una entrada distinta de cero, o *pivote*. Una buena alternativa es intercambiar las filas 1 y 4 (porque las comparaciones mentales en el siguiente paso no involucrarán fracciones).

Pivote

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

Columna pivote

Cree ceros debajo del pivote 1, para ello sume múltiplos de la primera fila a las filas de abajo, y obtenga la matriz (1) que se presenta enseguida. La posición pivote de la segunda fila debe estar lo más a la izquierda que sea posible —a saber, en la segunda columna—. Se elegirá al 2 en esta posición como el siguiente pivote.

Pivote

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \tag{1}$$

Próxima columna pivote

Sume la fila 2 multiplicado por $-5/2$ a la fila 3, y la fila 2 multiplicado por $3/2$ a la fila 4.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \tag{2}$$

La matriz en (2) es diferente a cualquiera de las matrices encontradas en la sección 1.1. ¡No hay forma de crear una entrada principal en la columna 3! (No pueden usarse las filas 1 o 2 porque al hacerlo se destruiría el arreglo escalonado de las entradas principales ya producidas.) Sin embargo, es posible producir una entrada principal en la columna 4 intercambiando las filas 3 y 4.

Pivote

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{Forma general:} \begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Columnas pivote

La matriz está en forma escalonada y, por lo tanto, las columnas 1, 2 y 4 de A son columnas pivote.

Posiciones pivote

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \tag{3}$$

Columnas pivote

Un **pivote**, como el ilustrado en el ejemplo 2, es un número distinto de cero situado en una posición pivote que se utiliza cuando es necesario para crear ceros por medio de operaciones de fila. Los pivotes empleados en el ejemplo 2 fueron 1, 2 y -5. Debe advertirse que estos números no son los mismos que los elementos reales de A ubicados en las posiciones pivote iluminadas que se muestran en (3). De hecho, una sucesión diferente de operaciones de fila podría involucrar un conjunto de pivotes distinto. Además, un pivote no será visible en la forma escalonada si la fila se escala para convertir el pivote en un 1 principal (lo cual muchas veces es conveniente para realizar cálculos a mano).

Con el ejemplo 2 como guía, ahora es posible describir un procedimiento eficiente para transformar una matriz en una matriz escalonada o escalonada reducida. El estudio cuidadoso y el dominio de este procedimiento producirán grandes dividendos durante todo el curso.

Algoritmo de reducción por filas

El algoritmo que se describe enseguida consta de cuatro pasos, y produce una matriz en forma escalonada. Un quinto paso produce una matriz en forma escalonada reducida. El algoritmo se ilustra mediante un ejemplo.

EJEMPLO 3 Aplique operaciones elementales de fila para transformar la siguiente matriz a la forma escalonada y después a la forma escalonada reducida:

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Solución

> **PASO 1**
> Empiece con la columna distinta de cero que se encuentra más a la izquierda. En este caso es una columna pivote. La posición pivote está en la parte superior.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ \underset{\uparrow\,\text{Columna pivote}}{3} & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

> **PASO 2**
> Seleccione como pivote una entrada distinta de cero en la columna pivote. Si es necesario, intercambie filas para mover esta entrada a la posición pivote.

Intercambie las filas 1 y 3. (También podrían haberse intercambiado las filas 1 y 2.)

$$\begin{bmatrix} \overset{\text{Pivote}}{3} & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

> **PASO 3**
> Use operaciones de reemplazo de fila para crear ceros en todas las posiciones ubicadas debajo del pivote.

Como paso preliminar, se podría dividir la fila superior entre el pivote, 3. Pero con dos números 3 en la columna 1, esto es tan fácil como sumar la fila 1 multiplicada por -1 a la fila 2.

$$\begin{bmatrix} \overset{\text{Pivote}}{3} & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

> **PASO 4**
> Cubra (o no tome en cuenta) la fila que contiene la posición pivote y cubra todas las filas, si existe alguna, por encima de ésta. Aplique los pasos 1, 2 y 3 a la submatriz restante. Repita el proceso hasta que no haya más filas distintas de cero por modificar.

Con la fila 1 cubierta, el paso 1 muestra que la columna 2 es la siguiente columna pivote; para el paso 2, en dicha columna se seleccionará como pivote la entrada "superior".

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & \overset{\text{Pivote}}{2} & -4 & 4 & 2 & -6 \\ 0 & \underset{\text{Nueva columna pivote}}{3} & -6 & 6 & 4 & -5 \end{bmatrix}$$

Para el paso 3, se podría insertar el paso opcional de dividir la fila “superior” de la submatriz entre el pivote 2. En vez de eso, se suma $-3/2$ veces la fila “superior” a la fila de abajo. Esto produce

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Cuando se cubre la fila que contiene la segunda posición pivote para el paso 4, queda una nueva submatriz que tiene solamente una fila:

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & \underset{\text{Pivote}}{1} & 4 \end{bmatrix}$$

Se ha alcanzado una forma escalonada para la matriz completa sin tener que aplicar los pasos 1, 2 y 3 en esta submatriz. Si se quisiera obtener la forma escalonada reducida, tendría que efectuarse un paso más.

> **PASO 5**
>
> Empiece con el pivote situado más a la derecha trabajando hacia arriba y a la izquierda, cree ceros arriba de cada pivote. Si un pivote no es 1, hágalo 1 mediante una operación de escalamiento.

El pivote situado más a la derecha está en la fila 3. Se crean ceros encima de él, sumando múltiplos adecuados de la fila 3 a las filas 2 y 1.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow \text{Fila } 1 + (-6)\cdot\text{Fila } 3 \\ \leftarrow \text{Fila } 2 + (-2)\cdot\text{Fila } 3 \\ \end{matrix}$$

El siguiente pivote está en la fila 2. Escale esta fila dividiéndola entre el pivote.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \\ \leftarrow \text{Fila escalada por } \tfrac{1}{2} \\ \end{matrix}$$

Se crea un cero en la columna 2 sumando 9 veces la fila 2 a la fila 1.

$$\begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow \text{Fila } 1 + (9)\cdot\text{Fila } 2 \\ \\ \end{matrix}$$

Por último, se escala la fila 1 al dividirla entre el pivote 3.

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad \leftarrow \text{Fila escalada por } \tfrac{1}{3}$$

Ésta es la forma escalonada reducida de la matriz original. ■

La combinación de los pasos 1 a 4 se llama fase progresiva del algoritmo de reducción por filas. El paso 5, que produce la forma escalonada reducida única, se llama **fase regresiva**.

NOTA NUMÉRICA

En el paso 2 que se mostró con anterioridad, un programa de computadora generalmente selecciona como pivote en una columna la entrada que tenga el mayor valor absoluto. Esta estrategia, llamada **pivoteo parcial**, se usa porque reduce los errores de redondeo en los cálculos.

Soluciones de sistemas lineales

El algoritmo de reducción por filas conduce directamente a una descripción explícita del conjunto solución de un sistema lineal cuando se aplica, el algoritmo, a la matriz aumentada del sistema.

Por ejemplo, suponga que la matriz aumentada de un sistema lineal ha sido transformada en la forma escalonada *reducida* equivalente

$$\begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Existen tres variables porque la matriz aumentada tiene cuatro columnas. El sistema de ecuaciones asociado es

$$\begin{aligned} x_1 \qquad - 5x_3 &= 1 \\ x_2 + \ \ x_3 &= 4 \\ 0 &= 0 \end{aligned} \tag{4}$$

Las variables x_1 y x_2 correspondientes a columnas pivote de la matriz se denominan **variables básicas**.[2] La otra variable, x_3, se llama **variable libre**.

Cuando un sistema es consistente, como en (4), el conjunto solución puede describirse de manera explícita al resolver el sistema de ecuaciones *reducido* para las variables básicas en términos de las variables libres. Esta operación es posible debido a que la

[2]Algunos textos utilizan el término *variables principales* porque corresponden a las columnas que contienen las entradas principales.

forma escalonada reducida coloca cada variable básica en una, y sólo una, ecuación. En (4), se puede despejar x_1 de la primera ecuación y x_2 de la segunda. (La tercera ecuación no se toma en cuenta porque no ofrece restricciones a las variables.)

$$\begin{cases} x_1 = 1 + 5x_3 \\ x_2 = 4 - x_3 \\ x_3 \text{ es libre} \end{cases} \tag{5}$$

Al afirmar que x_3 es "libre", se implica la posibilidad de asignarle cualquier valor. Una vez que se efectúa esta asignación, las fórmulas de (5) determinan los valores para x_1 y x_2. Por ejemplo, cuando $x_3 = 0$, la solución es (1, 4, 0); cuando $x_3 = 1$, la solución es (6, 3, 1). *Cada asignación diferente de x_3 determina una solución (diferente) del sistema, y cada solución del sistema está determinada por una asignación de x_3.*

La solución de (5) se denomina **solución general** del sistema porque proporciona una descripción explícita de *todas* las soluciones.

EJEMPLO 4 Encuentre la solución general del sistema lineal cuya matriz aumentada se ha reducido a

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

Solución La matriz está en forma escalonada, pero se requiere la forma escalonada reducida antes de despejar las variables básicas. A continuación se completa la reducción por filas. El símbolo ~ colocado antes de una matriz indica que ésta es equivalente por filas a la matriz precedente.

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 2 & -8 & 0 & 10 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

Existen cinco variables puesto que la matriz aumentada tiene seis columnas. Ahora el sistema asociado es

$$\begin{aligned} x_1 + 6x_2 \quad\; + 3x_4 \quad\; &= 0 \\ x_3 - 4x_4 \quad\; &= 5 \\ x_5 &= 7 \end{aligned} \tag{6}$$

Las columnas pivote de la matriz son 1, 3 y 5; así que las variables básicas son x_1, x_3 y x_5. Las variables restantes, x_2 y x_4, deben ser libres. Al despejar las variables básicas, se

obtiene la solución general:

$$\begin{cases} x_1 = -6x_2 - 3x_4 \\ x_2 \text{ es libre} \\ x_3 = 5 + 4x_4 \\ x_4 \text{ es libre} \\ x_5 = 7 \end{cases} \tag{7}$$

Observe que el valor de x_5 ya quedó fijado por la tercera ecuación del sistema (6).

Descripciones paramétricas de conjuntos solución

Las descripciones en (5) y (7) son *descripciones paramétricas* de conjuntos solución en los cuales las variables libres actúan como parámetros. *La resolución de un sistema* significa encontrar una descripción paramétrica del conjunto solución, o determinar que el conjunto solución está vacío.

Cuando un sistema es consistente y tiene variables libres, el conjunto solución permite obtener muchas descripciones paramétricas. Por ejemplo, en el sistema (4) se podría sumar cinco veces la ecuación 2 a la ecuación 1 y obtener el sistema equivalente

$$\begin{aligned} x_1 + 5x_2 \qquad &= 21 \\ x_2 + x_3 &= 4 \end{aligned}$$

Podría tratarse a x_2 como parámetro y despejar x_1 y x_3 en términos de x_2, y se tendría una descripción precisa del conjunto solución. Sin embargo, para ser consistente, se establece la convención (arbitraria) de usar siempre las variables libres como parámetros para describir un conjunto solución. (La sección de respuestas incluida al final del texto refleja también esta convención.)

Cuando un sistema es inconsistente, el conjunto solución está vacío, incluso si el sistema tiene variables libres. En este caso, el conjunto solución *no* tiene representación paramétrica.

Sustitución regresiva

Considere el sistema siguiente cuya matriz aumentada está en forma escalonada pero *no* en forma escalonada reducida:

$$\begin{aligned} x_1 - 7x_2 + 2x_3 - 5x_4 + 8x_5 &= 10 \\ x_2 - 3x_3 + 3x_4 + x_5 &= -5 \\ x_4 - x_5 &= 4 \end{aligned}$$

Un programa de computadora resolvería este sistema por sustitución regresiva, en lugar de calcular la forma escalonada reducida. Esto es, el programa resolvería la ecuación 3 para x_4 en términos de x_5 y sustituiría la expresión para x_4 en la ecuación 2; resolvería la ecuación 2 para x_2 y luego sustituiría las expresiones para x_2 y x_4 en la ecuación 1 y despejaría x_1.

El formato matricial que se utiliza en este texto para aplicar la fase regresiva de reducción por filas, la cual produce la forma escalonada reducida, requiere el mismo número de operaciones aritméticas que la sustitución regresiva. Pero la disciplina del formato matricial reduce sustancialmente la posibilidad de cometer errores durante los

cálculos efectuados a mano. Se recomienda de manera *enfática* usar solamente la forma escalonada *reducida* para resolver un sistema. La *Guía de estudio* (*Study Guide*) que acompaña a este texto ofrece algunas sugerencias útiles para realizar operaciones de fila con exactitud y rapidez.

NOTA NUMÉRICA

En general, la fase progresiva de la reducción por filas es mucho más larga que la fase regresiva. Para resolver un sistema, un algoritmo se mide generalmente en *flops* (u operaciones en punto flotante). Un **flop** es una operación aritmética ($+$, $-$, $*$, $/$) con dos números reales en punto flotante.[3] Para una matriz de $n \times (n + 1)$, la reducción a la forma escalonada puede requerir $2n^3/3 + n^2/2 - 7n/6$ flops (lo cual es aproximadamente $2n^3/3$ flops cuando n es moderadamente grande —por ejemplo, $n \geq 30$). Por otro lado, la reducción posterior a la forma escalonada reducida necesita cuando mucho n^2 flops.

Preguntas de existencia y unicidad

Aunque una forma escalonada no reducida es una herramienta poco eficiente para resolver un sistema, está considerada como el mecanismo correcto para resolver las dos preguntas fundamentales enunciadas en la sección 1.1.

EJEMPLO 5 Determine la existencia y unicidad de las soluciones del sistema

$$\begin{aligned} 3x_2 - 6x_3 + 6x_4 + 4x_5 &= -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15 \end{aligned}$$

Solución La matriz aumentada de este sistema se redujo por filas en el ejemplo 3 a

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \tag{8}$$

Las variables básicas son x_1, x_2 y x_5; las variables libres son x_3 y x_4. No hay ninguna ecuación del tipo $0 = 1$ que origine un sistema inconsistente, así que podría usarse sustitución regresiva para encontrar una solución. Pero en (8) ya es evidente la *existencia* de una solución. Además, la solución *no es única* porque existen variables libres. Cada asignación diferente de x_3 y x_4 determina una solución distinta. Por lo tanto, el sistema tiene un número infinito de soluciones.

[3]Tradicionalmente, un *flop* era sólo una multiplicación o una división porque la suma y la resta requerían mucho menos tiempo y podían no tomarse en cuenta. La definición de *flop* que se da aquí es la preferida en la actualidad, como consecuencia de los avances en la arquitectura de computadoras. Vea Golub y Van Loan, *Matrix Computations*, 2a. edición (Baltimore: The Johns Hopkins Press, 1989), pp. 19−20.

Cuando un sistema está en forma escalonada y no contiene ninguna ecuación del tipo $0 = b$, con b diferente de 0, toda ecuación distinta de cero contiene una variable básica con un coeficiente diferente de cero. Las variables básicas están completamente determinadas (sin variables libres), o por lo menos una de las variables básicas puede expresarse en términos de una o más variables libres. En el primer caso existe una solución única; en el último, hay un número infinito de soluciones (una para cada asignación de valores a las variables libres).

Estas observaciones justifican el teorema siguiente.

TEOREMA 2 Teorema de existencia y unicidad

Un sistema lineal es consistente si, y sólo si, la columna del extremo derecho de la matriz aumentada *no* es una columna pivote —esto es, si, y sólo si, una forma escalonada de la matriz aumentada *no* tiene ninguna fila de la forma

$$[0 \quad \cdots \quad 0 \quad b] \qquad \text{con } b \text{ diferente de cero.}$$

Si un sistema lineal es consistente, entonces el conjunto solución contiene (i) una solución única, cuando no existen variables libres, o bien (ii) un número infinito de soluciones, cuando existe por lo menos una variable libre.

El procedimiento siguiente define cómo encontrar y describir todas las soluciones de un sistema lineal.

USO DE LA REDUCCIÓN POR FILAS PARA RESOLVER UN SISTEMA LINEAL

1. Escriba la matriz aumentada del sistema.

2. Utilice el algoritmo de reducción por filas para obtener una matriz aumentada equivalente de forma escalonada. Decida si el sistema es o no consistente. Si no hay solución, deténgase; en caso contrario, continúe con el siguiente paso.

3. Continúe la reducción por filas hasta obtener la forma escalonada reducida.

4. Escriba el sistema de ecuaciones que corresponda a la matriz obtenida en el paso 3.

5. Reescriba cada ecuación diferente de cero del paso 4 de manera que su única variable básica esté expresada en términos de cualesquiera variables libres que aparezcan en la ecuación.

PROBLEMAS DE PRÁCTICA

1. Encuentre la solución general del sistema lineal cuya matriz aumentada es

$$\begin{bmatrix} 1 & -3 & -5 & 0 \\ 0 & 1 & 1 & 3 \end{bmatrix}$$

2. Encuentre la solución general del sistema

$$\begin{aligned} x_1 - 2x_2 - x_3 + 3x_4 &= 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 &= 3 \\ 3x_1 - 6x_2 - 6x_3 + 8x_4 &= 2 \end{aligned}$$

1.2 EJERCICIOS

En los ejercicios 1 y 2, determine cuáles matrices están en forma escalonada reducida y cuáles sólo en forma escalonada.

1. a. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$

2. a. $\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

d. $\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Reduzca por filas las matrices de los ejercicios 3 y 4 a la forma escalonada reducida. Encierre las posiciones pivote incluidas en la matriz final y en la matriz original, y enumere las columnas pivote.

3. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 7 \\ 6 & 7 & 8 & 9 \end{bmatrix}$ **4.** $\begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 1 \end{bmatrix}$

5. Describa las formas escalonadas posibles de una matriz de 2×2 distinta de cero. Utilice los símbolos (■), * y 0, como en la primera parte del ejemplo 1.

6. Repita el ejercicio 5 para una matriz de 3×2 diferente de cero.

Encuentre las soluciones generales de los sistemas cuyas matrices aumentadas se dan en los ejercicios 7 a 14.

7. $\begin{bmatrix} 1 & 3 & 4 & 7 \\ 3 & 9 & 7 & 6 \end{bmatrix}$ **8.** $\begin{bmatrix} 1 & 4 & 0 & 7 \\ 2 & 7 & 0 & 10 \end{bmatrix}$

9. $\begin{bmatrix} 0 & 1 & -6 & 5 \\ 1 & -2 & 7 & -6 \end{bmatrix}$ **10.** $\begin{bmatrix} 1 & -2 & -1 & 3 \\ 3 & -6 & -2 & 2 \end{bmatrix}$

11. $\begin{bmatrix} 3 & -4 & 2 & 0 \\ -9 & 12 & -6 & 0 \\ -6 & 8 & -4 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 1 & -7 & 0 & 6 & 5 \\ 0 & 0 & 1 & -2 & -3 \\ -1 & 7 & -4 & 2 & 7 \end{bmatrix}$

13. $\begin{bmatrix} 1 & -3 & 0 & -1 & 0 & -2 \\ 0 & 1 & 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & 1 & 9 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 2 & -5 & -6 & 0 & -5 \\ 0 & 1 & -6 & -3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 15 y 16 se utiliza la notación del ejemplo 1 para matrices en forma escalonada. Suponga que cada matriz representa la matriz aumentada para un sistema de ecuaciones lineales. En cada caso, determine si el sistema es consistente. De ser así, establezca si la solución es única.

15. a. $\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & 0 \end{bmatrix}$

b. $\begin{bmatrix} 0 & \blacksquare & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & \blacksquare \end{bmatrix}$

16. a. $\begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & 0 \end{bmatrix}$

b. $\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare & * \end{bmatrix}$

En los ejercicios 17 y 18, determine el valor o los valores de h tales que la matriz sea la matriz aumentada de un sistema lineal consistente.

17. $\begin{bmatrix} 2 & 3 & h \\ 4 & 6 & 7 \end{bmatrix}$

18. $\begin{bmatrix} 1 & -3 & -2 \\ 5 & h & -7 \end{bmatrix}$

En los ejercicios 19 y 20, elija h y k de tal forma que el sistema *a*) no tenga solución, *b*) tenga una solución única, y *c*) tenga muchas soluciones. Dé respuestas por separado para cada inciso.

19. $\begin{aligned} x_1 + hx_2 &= 2 \\ 4x_1 + 8x_2 &= k \end{aligned}$

20. $\begin{aligned} x_1 + 3x_2 &= 2 \\ 3x_1 + hx_2 &= k \end{aligned}$

En los ejercicios 21 y 22, señale cada enunciado como verdadero o falso. Justifique cada respuesta.[4]

21. a. En algunos casos, una matriz se puede reducir por filas a más de una matriz en forma escalonada reducida, usando diferentes secuencias de operaciones de fila.

b. El algoritmo de reducción por filas se aplica solamente a matrices aumentadas para un sistema lineal.

c. Una variable básica de un sistema lineal es una variable que corresponde a una columna pivote en la matriz de coeficientes.

d. Encontrar una descripción paramétrica del conjunto solución de un sistema lineal es lo mismo que *resolver* el sistema.

e. Si una fila en la forma escalonada de una matriz aumentada es $[0\ \ 0\ \ 0\ \ 5\ \ 0]$, entonces el sistema lineal asociado es inconsistente.

22. a. La forma escalonada de una matriz es única.

b. En una matriz, las posiciones pivote dependen de si se usan o no intercambios de fila en el proceso de reducción por filas.

c. La reducción de una matriz a forma escalonada se llama *fase progresiva* del proceso de reducción por filas.

d. Si un sistema tiene variables libres, el conjunto solución contiene muchas soluciones.

e. Una solución general de un sistema es una descripción explícita de todas las soluciones del sistema.

23. Suponga que una matriz de *coeficientes* de 3×5 para un sistema tiene tres columnas pivote. ¿Es consistente el sistema? ¿Por qué sí o por qué no?

24. Suponga que un sistema de ecuaciones lineales tiene una matriz *aumentada* de 3×5 cuya quinta columna es una columna pivote. ¿Es consistente el sistema? ¿Por qué sí o por qué no?

25. Suponga que la matriz de coeficientes de un sistema de ecuaciones lineales tiene una posición pivote en cada fila. Explique por qué este sistema es consistente.

26. Suponga que la matriz de coeficientes de un sistema lineal de tres ecuaciones en tres variables tiene un pivote en cada columna. Explique por qué tiene este sistema una solución única.

27. Reestructure la última oración del teorema 2 utilizando el concepto de columnas pivote: "Si un sistema lineal es consistente, entonces la solución es única si, y sólo si, __________ _____________."

28. ¿Qué debería saberse acerca de las columnas pivote de una matriz aumentada para advertir que el sistema lineal es consistente y tiene una solución única?

29. Un sistema de ecuaciones lineales con menos ecuaciones que incógnitas ocasionalmente se denomina *sistema subdeterminado.* Suponga que un sistema así resulta ser consistente. Explique por qué debería existir un número infinito de soluciones.

30. Proporcione el ejemplo de un sistema subdeterminado inconsistente de dos ecuaciones en tres incógnitas.

31. Un sistema de ecuaciones lineales con más ecuaciones que incógnitas ocasionalmente se denomina *sistema sobredeterminado.* ¿Puede ser consistente un sistema así? Ilustre su respuesta con un sistema específico de tres ecuaciones en dos incógnitas.

32. Suponga que una matriz de $n \times (n + 1)$ se reduce por filas a la forma escalonada reducida. Aproximadamente, ¿qué fracción del número total de operaciones (flops) está involucrada en la fase regresiva de la reducción cuando $n = 30$? ¿Cuándo $n = 300$?

Suponga que un conjunto de puntos en el plano representa datos experimentales. Un **polinomio de interpolación** para los datos es un polinomio cuya gráfica pasa por todos los puntos. En el trabajo científico, se puede usar un polinomio así, por ejemplo, para estimar valores entre los puntos de datos conocidos. Otro uso es crear curvas para imágenes gráficas en una pantalla de computadora. Un método apropiado para encontrar un polinomio de interpolación es resolver un sistema de ecuaciones lineales.

[4]Preguntas del tipo verdadero/falso como éstas aparecerán en muchas secciones. Los métodos para justificar sus respuestas se describieron antes de los ejercicios 23 y 24 de la sección 1.1.

33. Encuentre el polinomio de interpolación $p(t) = a_0 + a_1t + a_2t^2$ para los datos (1, 12), (2, 15), (3, 16). Esto es, encuentre a_0, a_1 y a_2 tales que

$$a_0 + a_1(1) + a_2(1)^2 = 12$$
$$a_0 + a_1(2) + a_2(2)^2 = 15$$
$$a_0 + a_1(3) + a_2(3)^2 = 16$$

34. **[M]** En un experimento de túnel de viento, la fuerza sobre un proyectil debida a la resistencia del aire se midió a diferentes velocidades:

Velocidad (100 pies/seg)	0	2	4	6	8	10
Fuerza (100 lb)	0	2.90	14.8	39.6	74.3	119

Encuentre un polinomio de interpolación para estos datos y estime la fuerza sobre el proyectil cuando éste viaja a 750 pies/seg. Utilice $p(t) = a_0 + a_1t + a_2t^2 + a_3t^3 + a_4t^4 + a_5t^5$. ¿Qué pasaría si se tratara de usar un polinomio con grado menor que 5? (Por ejemplo, pruebe con un polinomio cúbico.)[5]

[5] Los ejercicios marcados con el símbolo **[M]** están diseñados para resolverse con ayuda de un "Programa **M**atricial" (un programa de computadora, como **M**ATLAB, **M**aple, **M**athematica, **M**athCad o Derive, o una calculadora programable con capacidad para resolver matrices, como las calculadoras que fabrican Texas Instruments y Hewlett-Packard).

La solución general al sistema de ecuaciones es la línea de intersección de los dos planos.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. La forma escalonada reducida de la matriz aumentada y el sistema correspondiente son

$$\begin{bmatrix} 1 & 0 & -2 & 9 \\ 0 & 1 & 1 & 3 \end{bmatrix} \quad \text{y} \quad \begin{aligned} x_1 \qquad - 2x_3 &= 9 \\ x_2 + \quad x_3 &= 3 \end{aligned}$$

Las variables básicas son x_1 y x_2, y la solución general es

$$\begin{cases} x_1 = 9 + 2x_3 \\ x_2 = 3 - x_3 \\ x_3 \text{ es libre} \end{cases}$$

Nota: Resulta esencial que la solución general describa cada variable, con cualquier parámetro claramente identificado. El siguiente enunciado *no* describe la solución:

$$\begin{cases} x_1 = 9 + 2x_3 \\ x_2 = 3 - x_3 \\ x_3 = 3 - x_2 \quad \text{Solución incorrecta} \end{cases}$$

Esta descripción implica que *tanto* x_2 *como* x_3 son libres, lo cual desde luego no es el caso.

2. Al reducir por filas la matriz aumentada del sistema se obtiene:

$$\begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -6 & 8 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & -3 & -1 & 2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

Esta matriz escalonada muestra que el sistema es *inconsistente,* porque la columna de la extrema derecha es una columna pivote: la tercera fila corresponde a la ecuación $0 = 5$. No hay necesidad de realizar ninguna otra operación de fila. Observe que, en este problema, la presencia de las variables libres es irrelevante porque el sistema es inconsistente.

1.3 ECUACIONES VECTORIALES

Importantes propiedades de los sistemas lineales pueden ser descritas mediante el concepto y la notación de vectores. Esta sección relaciona ecuaciones que involucran vectores con sistemas de ecuaciones ordinarias. El término *vector* aparece en varios contextos matemáticos y físicos que se estudiarán en el capítulo 4, "Espacios vectoriales". Hasta entonces, el término *vector* se usará para denotar *una lista de números.* Esta idea sencilla permite realizar aplicaciones interesantes e importantes con la mayor rapidez posible.

Vectores en $\mathbb{R}^2$

Una matriz con una sola columna se llama **vector columna** o simplemente **vector**. Los siguientes son ejemplos de vectores con dos entradas

$$\mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} .2 \\ .3 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

donde w_1 y w_2 son cualesquiera números reales. El conjunto de todos los vectores con dos entradas se denota mediante $\mathbb{R}^2$ (lea "r-dos"). La $\mathbb{R}$ representa el conjunto de los números reales que aparecen como entradas en los vectores, y el exponente 2 indica que cada vector contiene dos entradas.[1]

Dos vectores en $\mathbb{R}^2$ son **iguales** si, y sólo si, sus entradas correspondientes son iguales. Así, $\begin{bmatrix} 4 \\ 7 \end{bmatrix}$ y $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$ *no* son iguales. Se dice que los vectores en $\mathbb{R}^2$ son *pares ordenados* de números reales.

Dados dos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^2$, su **suma** es el vector $\mathbf{u} + \mathbf{v}$ que se obtiene al sumar las entradas correspondientes de $\mathbf{u}$ y $\mathbf{v}$. Por ejemplo,

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1+2 \\ -2+5 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Dados un vector $\mathbf{u}$ y un número real c, el **múltiplo escalar** de $\mathbf{u}$ por c es el vector $c\mathbf{u}$ que se obtiene al multiplicar cada entrada de $\mathbf{u}$ por c. Por ejemplo,

$$\text{si } \mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{y} \quad c = 5, \qquad \text{entonces} \quad c\mathbf{u} = 5\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ -5 \end{bmatrix}$$

[1]La mayor parte del texto trata acerca de vectores y matrices que sólo tienen entradas reales. Sin embargo, todas las definiciones y teoremas de los capítulos 1 a 5, y de la mayor parte del texto restante, siguen siendo válidos cuando las entradas son números complejos. Los vectores y matrices complejos surgen de manera natural, por ejemplo, en ingeniería eléctrica y en física.

El número c de $c\mathbf{u}$ se llama **escalar**, y se escribe en letra cursiva para distinguirlo del vector en negritas $\mathbf{u}$.

Las operaciones de multiplicación por un escalar y suma de vectores se pueden combinar como en el siguiente ejemplo.

EJEMPLO 1 Dados $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, encuentre $4\mathbf{u}$, $(-3)\mathbf{v}$ y $4\mathbf{u} + (-3)\mathbf{v}$.

Solución

$$4\mathbf{u} = \begin{bmatrix} 4 \\ -8 \end{bmatrix}, \qquad (-3)\mathbf{v} = \begin{bmatrix} -6 \\ 15 \end{bmatrix}$$

y

$$4\mathbf{u} + (-3)\mathbf{v} = \begin{bmatrix} 4 \\ -8 \end{bmatrix} + \begin{bmatrix} -6 \\ 15 \end{bmatrix} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$$

Algunas veces, por conveniencia (y también para ahorrar espacio), se escribe un vector columna como $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ de la forma $(3, -1)$. En este caso, se usan paréntesis y una coma para distinguir el vector $(3, -1)$ de la matriz por filas 1×2 $[3, -1]$, que se escribe entre corchetes y sin coma. Así,

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} \neq [\, 3 \quad -1 \,]$$

porque las matrices tienen diferentes formas, aunque tengan las mismas entradas.

Descripciones geométricas de $\mathbb{R}^2$

Considere un sistema de coordenadas rectangulares en el plano. Como cada punto en el plano está determinado por un par ordenado de números, *puede identificarse un punto geométrico (a, b) con el vector columna* $\begin{bmatrix} a \\ b \end{bmatrix}$. Por lo tanto, puede considerarse a $\mathbb{R}^2$ como el conjunto de todos los puntos en el plano. Vea la figura 1.

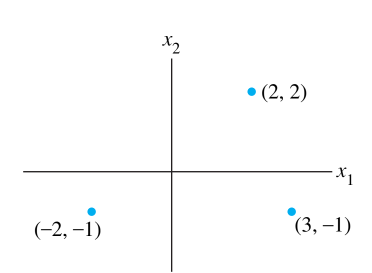


FIGURA 1
Vectores como puntos.

Con frecuencia, la visualización geométrica de un vector como $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ resulta beneficiada con la inclusión de una flecha (segmento de recta dirigido) desde el origen $(0, 0)$ hasta el punto $(3, -1)$, como en la figura 2. En este caso, los puntos individuales a lo largo de la flecha no tienen significado especial.[2]

La suma de dos vectores tiene una representación geométrica útil. La siguiente regla puede verificarse por medio de geometría analítica.

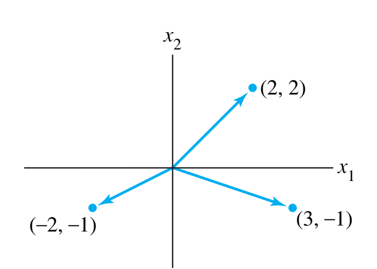


FIGURA 2
Vectores con flechas.

[2]En física, las flechas pueden representar fuerzas y, por lo general, son libres de moverse en el espacio. Esta interpretación de los vectores se estudiará en la sección 4.1.

REGLA DEL PARALELOGRAMO PARA LA SUMA

Si **u** y **v** en $\mathbb{R}^2$ se representan como puntos en el plano, entonces **u** + **v** corresponde al cuarto vértice del paralelogramo cuyos otros vértices son **u**, **0** y **v**. Vea la figura 3.

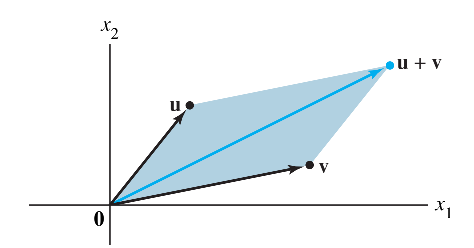


FIGURA 3 La regla del paralelogramo.

EJEMPLO 2 Los vectores $\mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$, y $\mathbf{u} + \mathbf{v} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$ se representan en la figura 4.

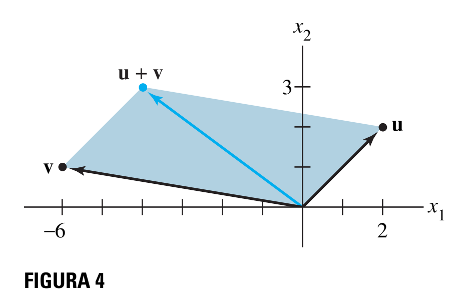


FIGURA 4

El siguiente ejemplo ilustra el hecho de que el conjunto de todos los múltiplos escalares de un vector fijo es una recta que pasa por el origen, (0, 0).

EJEMPLO 3 Sea $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. Represente en una gráfica los vectores **u**, 2**u** y $-\frac{2}{3}\mathbf{u}$.

Solución Vea la figura 5, donde se muestran **u**, $2\mathbf{u} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$, y $-\frac{2}{3}\mathbf{u} = \begin{bmatrix} -2 \\ 2/3 \end{bmatrix}$. La flecha para 2**u** tiene el doble de largo que la empleada para **u**, y ambas apuntan en la misma dirección. La flecha para $-\frac{2}{3}\mathbf{u}$ es dos tercios del largo de la flecha para **u**, y las dos apuntan en direcciones opuestas. En general, la longitud de la flecha para $c\mathbf{u}$ es $|c|$ veces la

longitud de la flecha para **u**. [Recuerde que la longitud del segmento de recta desde (0, 0) hasta (a, b) es $\sqrt{a^2 + b^2}$. Esto se analizará más a fondo en el capítulo 6.]

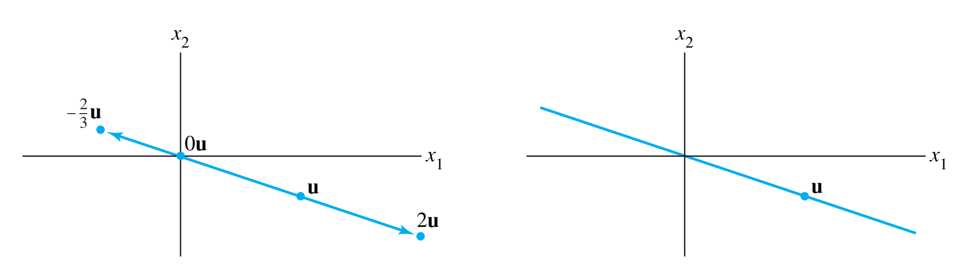


FIGURA 5

Vectores en $\mathbb{R}^3$

Los vectores en $\mathbb{R}^3$ son matrices columna de 3×1 con tres entradas. Se representan geométricamente por medio de puntos en un espacio coordenado de tres dimensiones, algunas veces se incluyen flechas desde el origen para proporcionar mayor claridad visual. Los vectores $\mathbf{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ y $2\mathbf{a}$ se muestran en la figura 6.

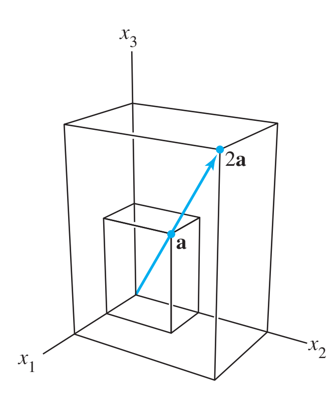


FIGURA 6
Múltiplos escalares en $\mathbb{R}^3$.

Vectores en $\mathbb{R}^n$

Si n es un entero positivo, $\mathbb{R}^n$ (lea "r-n") denota la colección de todas las listas (o *n-adas ordenadas)* de n números reales, escritas, por lo general, como matrices columna de $n \times 1$ del tipo

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

El vector cuyas entradas son todas iguales a cero se llama **vector cero** y se denota mediante $\mathbf{0}$. (El número de entradas en $\mathbf{0}$ será evidente a partir del contexto.)

La igualdad de vectores en $\mathbb{R}^n$ y las operaciones de multiplicación escalar y suma de vectores en $\mathbb{R}^n$ se definen entrada por entrada igual que en $\mathbb{R}^2$. Estas operaciones de vectores tienen las siguientes propiedades, que se pueden verificar en forma directa a partir de las propiedades correspondientes para números reales. Vea el problema de práctica 1 y los ejercicios 33 y 34 incluidos al final de esta sección.

PROPIEDADES ALGEBRAICAS DE $\mathbb{R}^n$

Para todos **u**, **v** y **w** en $\mathbb{R}^n$ y todos los escalares c y d:

(i) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$

(ii) $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$

(iii) $\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$

(iv) $\mathbf{u}+(-\mathbf{u})=-\mathbf{u}+\mathbf{u}=\mathbf{0}$, donde $-\mathbf{u}$ denota a $(-1)\mathbf{u}$

(v) $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$

(vi) $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$

(vii) $c(d\mathbf{u})=(cd)(\mathbf{u})$

(viii) $1\mathbf{u}=\mathbf{u}$.

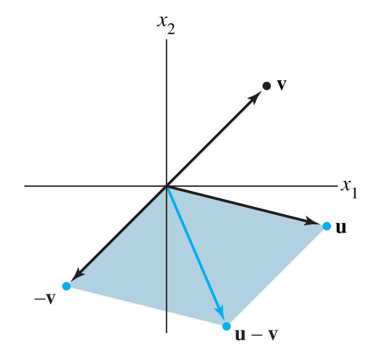


FIGURA 7
Resta de vectores.

Para simplificar la notación, también se utiliza "resta de vectores" y se escribe $\mathbf{u}-\mathbf{v}$ en lugar de $\mathbf{u}+(-1)\mathbf{v}$. En la figura 7 se muestra a $\mathbf{u}-\mathbf{v}$ como la suma de **u** y $-\mathbf{v}$.

Combinaciones lineales

Dados los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ en $\mathbb{R}^n$ y los escalares $c_1, c_2, \ldots, c_p$, el vector **y** definido por

$$\mathbf{y}=c_1\mathbf{v}_1+\cdots+c_p\mathbf{v}_p$$

se llama **combinación lineal** de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ con **pesos** $c_1, c_2, \ldots, c_p$. La propiedad (ii) enunciada anteriormente permite omitir los paréntesis cuando se forma una combinación lineal de este tipo. En una combinación lineal, los pesos pueden ser cualesquiera números reales, incluso el cero. Por ejemplo, algunas combinaciones lineales de los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ son

$$\sqrt{3}\mathbf{v}_1+\mathbf{v}_2, \qquad \tfrac{1}{2}\mathbf{v}_1\ (=\tfrac{1}{2}\mathbf{v}_1+0\mathbf{v}_2), \qquad \text{y} \qquad \mathbf{0}\ (=0\mathbf{v}_1+0\mathbf{v}_2)$$

EJEMPLO 4 En la figura 8 se identifican algunas combinaciones lineales seleccionadas de $\mathbf{v}_1=\begin{bmatrix}-1\\1\end{bmatrix}$ y $\mathbf{v}_2=\begin{bmatrix}2\\1\end{bmatrix}$. (Observe que los conjuntos de líneas paralelas de la rejilla están trazados mediante múltiplos enteros de $\mathbf{v}_1$ y $\mathbf{v}_2$.) Estime las combinaciones lineales de $\mathbf{v}_1$ y $\mathbf{v}_2$ que generan los vectores **u** y **w**.

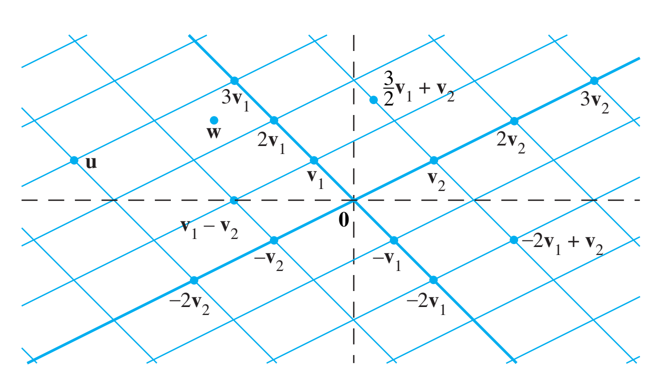


FIGURA 8 Combinaciones lineales de $\mathbf{v}_1$ y $\mathbf{v}_2$.

Solución La regla del paralelogramo muestra que $\mathbf{u}$ es la suma de $3\mathbf{v}_1$ y $-2\mathbf{v}_2$, esto es,

$$\mathbf{u} = 3\mathbf{v}_1 - 2\mathbf{v}_2$$

Esta expresión para $\mathbf{u}$ puede interpretarse como las instrucciones para viajar desde el origen hasta $\mathbf{u}$ a lo largo de dos rutas rectas. Primero, viaje tres unidades en la dirección $\mathbf{v}_1$ hasta $3\mathbf{v}_1$, y después viaje -2 unidades en la dirección $\mathbf{v}_2$ (paralela a la línea que pasa por $\mathbf{v}_2$ y $\mathbf{0}$). Enseguida, aunque el vector $\mathbf{w}$ no está en una línea de la rejilla, parece que $\mathbf{w}$ queda aproximadamente a media distancia entre dos pares de líneas de la rejilla, en el vértice de un paralelogramo determinado por $(5/2)\mathbf{v}_1$ y $(-1/2)\mathbf{v}_2$. (Vea la figura 9.) Así,

$$\mathbf{w} = \tfrac{5}{2}\mathbf{v}_1 - \tfrac{1}{2}\mathbf{v}_2$$

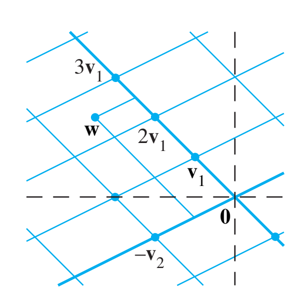


FIGURA 9

El siguiente ejemplo relaciona un problema de combinaciones lineales con la pregunta fundamental de existencia que se estudió en las secciones 1.1 y 1.2.

EJEMPLO 5 Sean $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$, y $\mathbf{b} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$. Determine si $\mathbf{b}$ puede generarse (o escribirse) como una combinación lineal de $\mathbf{a}_1$ y $\mathbf{a}_2$. Esto es, calcule si existen pesos x_1 y x_2 tales que

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b} \tag{1}$$

Si la ecuación vectorial (1) tiene solución, encuéntrela.

Solución Utilice las definiciones de multiplicación escalar y suma de vectores para reescribir la ecuación vectorial

$$x_1\underset{\uparrow\atop \mathbf{a}_1}{\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}} + x_2\underset{\uparrow\atop \mathbf{a}_2}{\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}} = \underset{\uparrow\atop \mathbf{b}}{\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}}$$

la cual es la misma que

$$\begin{bmatrix} x_1 \\ -2x_1 \\ -5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

y

$$\begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{2}$$

Los vectores que aparecen en los lados derecho e izquierdo de (2) son iguales si, y sólo si, sus entradas correspondientes son iguales. Esto es, x_1 y x_2 hacen que la ecuación

vectorial (1) sea verdadera si, y sólo si, x_1 y x_2 satisfacen el sistema

$$\begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= -3 \end{aligned} \tag{3}$$

Este sistema se resuelve al reducir por filas su matriz aumentada, como se muestra a continuación:[3]

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

La solución de (3) es $x_1 = 3$ y $x_2 = 2$. Por lo tanto, **b** es una combinación lineal de $\mathbf{a}_1$ y $\mathbf{a}_2$, con pesos $x_1 = 3$ y $x_2 = 2$. Esto es,

$$3\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

Observe en el ejemplo 5 que los vectores originales $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{b}$ son las columnas de la matriz aumentada que se redujo por filas:

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix}$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{b} \end{matrix}$$

Ahora se escribirá esta matriz de tal modo que llame la atención hacia sus columnas a saber,

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{b}] \tag{4}$$

Queda claro cómo, sin seguir los pasos intermedios del ejemplo 5, puede escribirse la matriz aumentada directamente a partir de la ecuación vectorial (1). Simplemente se toman los vectores en el orden en que aparecen en (1) y se colocan en las columnas de una matriz como en (4).

El análisis anterior se puede modificar con facilidad para establecer el siguiente hecho fundamental.

Una ecuación vectorial como

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

tiene el mismo conjunto solución que el sistema lineal cuya matriz aumentada es:

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \quad \mathbf{b}] \tag{5}$$

En particular, **b** se puede generar mediante una combinación lineal de $\mathbf{a}_1, \ldots, \mathbf{a}_n$ si, y sólo si, existe una solución del sistema lineal que corresponde a (5).

[3]El símbolo $\sim$ colocado entre las matrices denota equivalencia por filas (sección 1.2).

Una de las ideas fundamentales del álgebra lineal es estudiar el conjunto de todos los vectores que puedan generarse o escribirse como una combinación lineal de un conjunto de vectores fijo $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$.

DEFINICIÓN

Si $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en $\mathbb{R}^n$, entonces el conjunto de todas las combinaciones linea-les de $\mathbf{v}_1, \ldots, \mathbf{v}_p$ se denota mediante Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ y recibe el nombre de **subespacio de** $\mathbb{R}^n$ **generado por** $\mathbf{v}_1, \ldots, \mathbf{v}_p$. Esto es, Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es la colección de todos los vectores que pueden escribirse en la forma

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

donde $c_1, \ldots, c_p$ son escalares.

Preguntar si un vector $\mathbf{b}$ está en Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ equivale a preguntar si la ecuación vectorial

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{b}$$

tiene una solución o, de manera equivalente, si el sistema lineal con matriz aumentada $[\mathbf{v}_1 \quad \cdots \mathbf{v}_p \quad \mathbf{b}]$ tiene una solución.

Advierta que Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ contiene todos los múltiplos escalares de $\mathbf{v}_1$ (por ejemplo), puesto que $c\mathbf{v}_1 = c\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. En particular, el vector cero debe estar en Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

Una descripción geométrica de Gen{v} y Gen{u, v}

Sea $\mathbf{v}$ un vector diferente de cero en $\mathbb{R}^3$. Entonces Gen$\{\mathbf{v}\}$ es el conjunto de todos los múltiplos escalares de $\mathbf{v}$, y se visualiza como el conjunto de puntos sobre la línea en $\mathbb{R}^3$ que pasa por $\mathbf{v}$ y $\mathbf{0}$. Vea la figura 10.

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores diferentes de cero en $\mathbb{R}^3$, y $\mathbf{v}$ no es un múltiplo de $\mathbf{u}$, entonces Gen$\{\mathbf{u}, \mathbf{v}\}$ es el plano en $\mathbb{R}^3$ que contiene a $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{0}$. En particular, Gen$\{\mathbf{u}, \mathbf{v}\}$ contiene la línea en $\mathbb{R}^3$ que pasa por $\mathbf{u}$ y $\mathbf{0}$ y la línea que pasa por $\mathbf{v}$ y $\mathbf{0}$. Vea la figura 11.

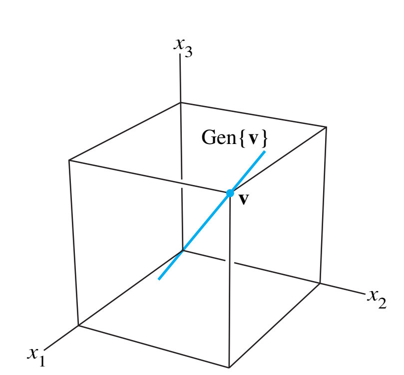


FIGURA 10 Gen$\{\mathbf{v}\}$ como una línea que pasa por el origen.

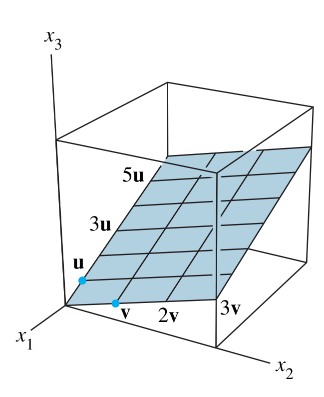


FIGURA 11 Gen$\{\mathbf{u}, \mathbf{v}\}$ como un plano que pasa por el origen.

EJEMPLO 6 Sea $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$, y $\mathbf{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$. Entonces Gen$\{\mathbf{a}_1, \mathbf{a}_2\}$ es un plano que pasa por el origen en $\mathbb{R}^3$. ¿Está $\mathbf{b}$ en ese plano?

Solución ¿Tiene solución la ecuación $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$? Para contestar esto, reduzca por filas la matriz aumentada $[\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{b}]$:

$$\begin{bmatrix} 1 & 5 & -3 \\ -2 & -13 & 8 \\ 3 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & -18 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{bmatrix}$$

La tercera ecuación es $0x_2 = -2$, lo cual muestra que el sistema no tiene solución. La ecuación vectorial $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$ no tiene solución y, por lo tanto, $\mathbf{b}$ *no* está en Gen$\{\mathbf{a}_1, \mathbf{a}_2\}$.

Combinaciones lineales en aplicaciones

El ejemplo final muestra cómo pueden surgir múltiplos escalares y combinaciones lineales cuando una cantidad, tal como un “costo”, se descompone en varias categorías. El principio básico para el ejemplo se relaciona con el costo de producir varias unidades de un artículo cuando se conoce el costo por unidad:

$$\left\{\begin{matrix}\text{número de} \\ \text{unidades}\end{matrix}\right\} \cdot \left\{\begin{matrix}\text{costo por} \\ \text{unidad}\end{matrix}\right\} = \left\{\begin{matrix}\text{costo} \\ \text{total}\end{matrix}\right\}$$

EJEMPLO 7 Una compañía fabrica dos productos. Para \$1.00 obtenido del producto B, la compañía gasta \$.45 en materiales, \$.25 en mano de obra, y \$.15 en gastos generales. Para \$1.00 obtenido del producto C, la compañía gasta \$.40 en materiales, \$.30 en mano de obra, y \$.15 en gastos generales. Sean

$$\mathbf{b} = \begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} \quad \text{y} \quad \mathbf{c} = \begin{bmatrix} .40 \\ .30 \\ .15 \end{bmatrix}$$

entonces $\mathbf{b}$ y $\mathbf{c}$ representan los “costos por dólar de ingreso” de los dos productos.

a. ¿Qué interpretación económica puede darse al vector 100$\mathbf{b}$?

b. Suponga que la compañía desea fabricar x_1 dólares del producto B y x_2 dólares del producto C. Proporcione un vector que describa los diversos costos que tendrá esta empresa (por materiales, mano de obra y gastos generales).

Solución

a. Se tiene

$$100\mathbf{b} = 100\begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} = \begin{bmatrix} 45 \\ 25 \\ 15 \end{bmatrix}$$

El vector l00$\mathbf{b}$ enlista los diversos costos por generar \$100 del producto B —a saber, \$45 por materiales, \$25 por mano de obra, y \$15 por gastos generales.

b. Los costos de obtener x_1 dólares a partir de B están dados por el vector $x_1\mathbf{b}$, y los costos de obtener x_2 dólares del producto C están dados por $x_2\mathbf{c}$. Por lo tanto, el costo total de ambos productos lo proporciona el vector $x_1\mathbf{b} + x_2\mathbf{c}$.

PROBLEMAS DE PRÁCTICA

1. Demuestre que $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ para todos $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$.

2. Determine los valores de h para los que $\mathbf{y}$ estará en Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

1.3 EJERCICIOS

En los ejercicios 1 y 2, calcule $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - 2\mathbf{v}$.

1. $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$

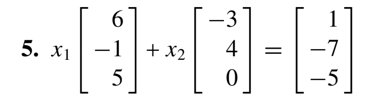
2. $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

En los ejercicios 3 y 4, represente los siguientes vectores utilizando flechas en una gráfica xy: $\mathbf{u}$, $\mathbf{v}$, $-\mathbf{v}$, $-2\mathbf{v}$, $\mathbf{u} + \mathbf{v}$, $\mathbf{u} - \mathbf{v}$ y $\mathbf{u} - 2\mathbf{v}$. Observe que $\mathbf{u} - \mathbf{v}$ es el vértice de un paralelogramo cuyos otros vértices son $\mathbf{u}$, $\mathbf{0}$ y $-\mathbf{v}$.

3. $\mathbf{u}$ y $\mathbf{v}$ como en el ejercicio 1.

4. $\mathbf{u}$ y $\mathbf{v}$ como en el ejercicio 2.

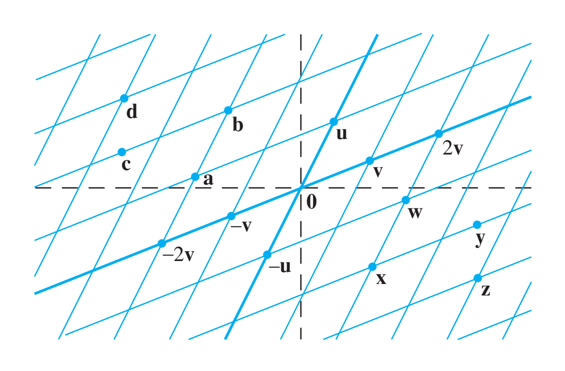
En los ejercicios 5 y 6, escriba un sistema de ecuaciones que sea equivalente a la ecuación vectorial dada.



5. $x_1 \begin{bmatrix} 6 \\ -1 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ -5 \end{bmatrix}$

6. $x_1 \begin{bmatrix} -2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 8 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Use la siguiente figura para escribir cada vector enlistado en los ejercicios 7 y 8 como una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$. ¿Cada vector en $\mathbb{R}^2$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$?



7. Los vectores $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ y $\mathbf{d}$.

8. Los vectores $\mathbf{w}$, $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$.

En los ejercicios 9 y 10, escriba una ecuación vectorial que sea equivalente al sistema de ecuaciones dado.

9.
$$\begin{aligned} x_2 + 5x_3 &= 0 \\ 4x_1 + 6x_2 - x_3 &= 0 \\ -x_1 + 3x_2 - 8x_3 &= 0 \end{aligned}$$

10.
$$\begin{aligned} 4x_1 + x_2 + 3x_3 &= 9 \\ x_1 - 7x_2 - 2x_3 &= 2 \\ 8x_1 + 6x_2 - 5x_3 &= 15 \end{aligned}$$

En los ejercicios 11 y 12, determine si $\mathbf{b}$ es una combinación lineal de $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$.

11. $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 5 \\ -6 \\ 8 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$

12. $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 5 \\ 5 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -5 \\ 11 \\ -7 \end{bmatrix}$

En los ejercicios 13 y 14, determine si **b** es una combinación lineal de los vectores formados a partir de las columnas de la matriz A.

13. $A = \begin{bmatrix} 1 & -4 & 2 \\ 0 & 3 & 5 \\ -2 & 8 & -4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ -7 \\ -3 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & -2 & -6 \\ 0 & 3 & 7 \\ 1 & -2 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 11 \\ -5 \\ 9 \end{bmatrix}$

En los ejercicios 15 y 16, enliste cinco vectores incluidos en Gen[$\mathbf{v}_1$, $\mathbf{v}_2$]. Para cada vector, muestre los pesos usados en $\mathbf{v}_1$ y $\mathbf{v}_2$ para generar el vector y enliste las tres entradas del vector. No haga ningún bosquejo.

15. $\mathbf{v}_1 = \begin{bmatrix} 7 \\ 1 \\ -6 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix}$

16. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 3 \end{bmatrix}$

17. Sean $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} -2 \\ -3 \\ 7 \end{bmatrix}$, y $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ h \end{bmatrix}$. ¿Para cuáles valores de h está **b** en el plano generado por $\mathbf{a}_1$ y $\mathbf{a}_2$?

18. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -3 \\ 1 \\ 8 \end{bmatrix}$, y $\mathbf{y} = \begin{bmatrix} h \\ -5 \\ -3 \end{bmatrix}$. ¿Para cuáles valores de h está **b** en el plano generado por $\mathbf{v}_1$ y $\mathbf{v}_2$?

19. Dé una descripción geométrica de Gen{$\mathbf{v}_1$, $\mathbf{v}_2$} para los vectores

$$\mathbf{v}_1 = \begin{bmatrix} 8 \\ 2 \\ -6 \end{bmatrix} \text{ y } \mathbf{v}_2 = \begin{bmatrix} 12 \\ 3 \\ -9 \end{bmatrix}.$$

20. Dé una descripción geométrica de Gen{$\mathbf{v}_1$, $\mathbf{v}_2$} para los vectores del ejercicio 16.

21. Sean $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Muestre que $\begin{bmatrix} h \\ k \end{bmatrix}$ está en Gen{**u**, **v**} para todas h y k.

22. Construya una matriz A de 3 × 3, con entradas distintas de cero, y un vector **b** en $\mathbb{R}^3$ tal que **b** *no* esté en el conjunto generado por las columnas de A.

En los ejercicios 23 y 24, señale cada enunciado como verdadero o falso. Justifique cada una de sus respuestas.

23. a. Una notación distinta para el vector $\begin{bmatrix} -4 \\ 3 \end{bmatrix}$ es $[-4 \quad 3]$.

b. Los puntos en el plano correspondientes a $\begin{bmatrix} -2 \\ 5 \end{bmatrix}$ y $\begin{bmatrix} -5 \\ 2 \end{bmatrix}$ están sobre una línea que pasa por el origen.

c. Un ejemplo de una combinación lineal de los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ es el vector $\frac{1}{2}\mathbf{v}_1$.

d. El conjunto solución del sistema lineal cuya matriz aumentada es $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \quad \mathbf{b}]$ es igual al conjunto solución de la ecuación $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}$.

e. El conjunto Gen{**u**, **v**} siempre se visualiza como un plano que pasa por el origen.

24. a. Cualquier lista de cinco números reales es un vector en $\mathbb{R}^3$.

b. El vector **u** resulta cuando al vector **u** – **v** se le suma el vector **v**.

c. Los pesos $c_1, \ldots, c_p$ en una combinación lineal $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$ no pueden ser todos iguales a cero.

d. Cuando **u** y **v** son vectores distintos de cero, Gen{**u**, **v**} contiene la línea que pasa por **u** y por el origen.

e. Preguntar si el sistema lineal correspondiente a una matriz aumentada $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \quad \mathbf{b}]$ tiene una solución es lo mismo que preguntar si **b** está en Gen[$\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$].

25. Sean $A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 3 & -2 \\ -2 & 6 & 3 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ -4 \end{bmatrix}$. Denote las columnas de A mediante $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, y sea $W =$ Gen{$\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$}.

a. ¿Está **b** en {$\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$}? ¿Cuántos vectores están en {$\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$}?

b. ¿Está **b** en W? ¿Cuántos vectores están en W?

c. Muestre que $\mathbf{a}_1$ está en W. [*Indicación:* No se requieren operaciones de fila.]

26. Sea $A = \begin{bmatrix} 2 & 0 & 6 \\ -1 & 8 & 5 \\ 1 & -2 & 1 \end{bmatrix}$, sea $\mathbf{b} = \begin{bmatrix} 10 \\ 3 \\ 3 \end{bmatrix}$, y sea W el conjunto de todas las combinaciones lineales de las columnas de A.

a. ¿Está **b** en W?

b. Muestre que la tercera columna de A está en W.

27. Una compañía minera tiene dos minas. Las operaciones de un día en la mina 1 producen mineral que contiene 20 toneladas métricas de cobre y 550 kilogramos (kg) de plata,

mientras que las operaciones de un día en la mina 2 producen mineral que contiene 30 toneladas métricas de cobre y 500 kg de plata. Sean $\mathbf{v}_1 = \begin{bmatrix} 20 \\ 550 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 30 \\ 500 \end{bmatrix}$. Entonces, $\mathbf{v}_1$ y $\mathbf{v}_2$ representan el "rendimiento diario" de las minas 1 y 2, respectivamente.

a. ¿Que interpretación física puede dársele al vector $5\mathbf{v}_1$?

b. Suponga que la compañía trabaja la mina l durante x_1 días y la mina 2 x_2 días. Escriba una ecuación vectorial cuya solución proporcione el número de días que deba trabajarse cada mina para producir 150 toneladas de cobre y 2825 kg de plata. No resuelva la ecuación.

c. **[M]** Resuelva la ecuación del inciso (*b*).

28. Una planta de vapor quema dos clases de carbón: antracita (A) y bituminoso (B). Por cada tonelada de A quemada, la planta produce 27.6 millones de Btu de calor, 3100 gramos (g) de dióxido de azufre, y 250 g de materia en partículas (de sólidos contaminantes). Por cada tonelada de B quemada, se producen 30.2 millones de Btu, 6400 g de dióxido de azufre, y 360 g de materia en partículas.

a. ¿Cuánto calor produce la planta de vapor cuando quema x_1 toneladas de A y x_2 toneladas de B?

b. Suponga que el rendimiento de la planta de vapor se describe con un vector que enlista las cantidades de calor, dióxido de azufre y materia en partículas. Exprese este rendimiento como una combinación lineal de dos vectores, para ello suponga que la planta quema x_1 toneladas de A y x_2 toneladas de B.

c. **[M]** Durante cierto periodo, la planta de vapor produce 162 millones de Btu de calor, 23,610 g de dióxido de azufre, y 1623 g de materia en partículas. Determine cuántas toneladas de cada tipo de carbón debe quemar esta planta. Incluya una ecuación vectorial como parte de su solución.

29. Sean $\mathbf{v}_1, \ldots, \mathbf{v}_k$ puntos en $\mathbb{R}^3$, y suponga que para $j = 1, \ldots, k$ un objeto con masa m_j se localiza en el punto $\mathbf{v}_j$. Los físicos llaman a tales objetos *masas puntuales.* La masa total del sistema de masas puntuales es

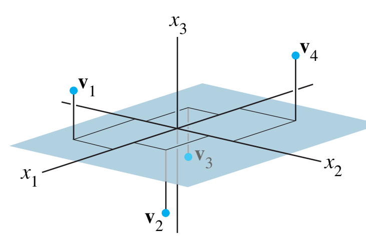
$$m = m_1 + \cdots + m_k$$

El centro de gravedad (o *centro de masa)* del sistema es

$$\bar{\mathbf{v}} = \frac{1}{m}[m_1\mathbf{v}_1 + \cdots + m_k\mathbf{v}_k]$$

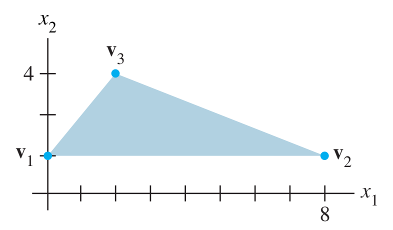
Calcule el centro de gravedad del sistema constituido por las siguientes masas puntuales (vea la figura):

Punto	Masa
$\mathbf{v}_1 = (5, -4, 3)$	2 g
$\mathbf{v}_2 = (4, 3, -2)$	5 g
$\mathbf{v}_3 = (-4, -3, -1)$	2 g
$\mathbf{v}_4 = (-9, 8, 6)$	1 g



30. Sea $\mathbf{v}$ el centro de masa de un sistema de masas puntuales localizado en $\mathbf{v}_1, \ldots, \mathbf{v}_k$ como en el ejercicio 29. ¿Está $\mathbf{v}$ en Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$? Explique su respuesta.

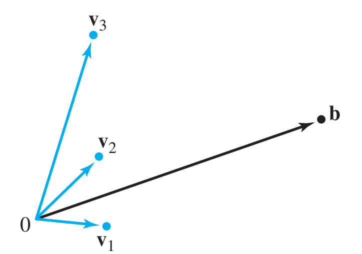
31. La placa triangular delgada que se muestra en la siguiente figura tiene espesor y densidad uniformes con vértices en $\mathbf{v}_1 = (0, 1)$, $\mathbf{v}_2 = (8, 1)$, y $\mathbf{v}_3 = (2, 4)$, y su masa es de 3 g.



a. Encuentre las coordenadas (x, y) del centro de masa de la placa. Este "punto de balance" coincide con el centro de masa de un sistema que consta de tres masas puntuales ubicadas en los vértices de la placa.

b. Determine cómo distribuir una masa adicional de 6 g en los tres vértices de la placa para trasladar el punto de balance a (2, 2). [*Indicación:* Sean w_1, w_2 y w_3 las masas agregadas a los tres vértices, de manera que $w_1 + w_2 + w_3 = 6$.]

32. Considere los vectores $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{b}$ en $\mathbb{R}^2$ que se muestran en la figura. ¿La ecuación $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}$ tiene alguna solución? ¿La solución es única? Utilice la figura para explicar sus respuestas.



33. Use los vectores $\mathbf{u} = (u_1, \ldots, u_n)$, $\mathbf{v} = (v_1, \ldots, v_n)$, y $\mathbf{w} = (w_1, \ldots, w_n)$ para verificar las siguientes propiedades algebraicas de $\mathbb{R}^n$.

a. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

b. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ para cada escalar c

34. Use el vector $\mathbf{u} = (u_1, \ldots, u_n)$ para verificar las siguientes propiedades algebraicas de $\mathbb{R}^n$.

a. $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$

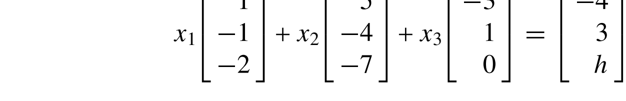
b. $c(d\mathbf{u}) = (cd)\mathbf{u}$ para todos los escalares c y d

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

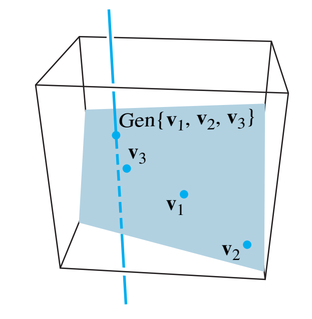
1. Tome vectores arbitrarios $\mathbf{u} = (u_1, \ldots, u_n)$ y $\mathbf{v} = (v_1, \ldots, v_n)$ en $\mathbb{R}^n$, y calcule

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1 + v_1, \ldots, u_n + v_n) && \text{Definición de la suma de vectores} \\ &= (v_1 + u_1, \ldots, v_n + u_n) && \text{Conmutatividad de la suma en } \mathbb{R} \\ &= \mathbf{v} + \mathbf{u} && \text{Definición de la suma de vectores} \end{aligned}$$

2. El vector $\mathbf{y}$ pertenece a Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si, y sólo si, existen escalares x_1, x_2, x_3 tales que



$$x_1\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} + x_2\begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix} + x_3\begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$



Los puntos $\begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$ están sobre una línea que interseca el plano cuando $h = 5$.

Esta ecuación vectorial es equivalente a un sistema de tres ecuaciones lineales con tres incógnitas. Si se reduce por filas la matriz aumentada de este sistema, se tiene que

$$\begin{bmatrix} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ -2 & -7 & 0 & h \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 3 & -6 & h-8 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{bmatrix}$$

El sistema es consistente si, y sólo si, no hay pivote en la cuarta columna. Esto es, $h - 5$ debe ser 0. Así que $\mathbf{y}$ está en Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si, y sólo si, $h = 5$.

Recuerde: La presencia de una variable libre en un sistema no garantiza que el sistema sea consistente.

1.4 LA ECUACIÓN MATRICIAL $A\mathbf{x} = \mathbf{b}$

Una idea fundamental en el álgebra lineal es visualizar una combinación lineal de vectores como el producto de una matriz y un vector. La siguiente definición permitirá expresar de otra manera algunos de los conceptos de la sección 1.3.

DEFINICIÓN

Si A es una matriz de $m \times n$, con columnas $\mathbf{a}_1, \ldots, \mathbf{a}_n$, y si $\mathbf{x}$ está en $\mathbb{R}^n$, entonces el **producto de** A **y** $\mathbf{x}$, denotado por $A\mathbf{x}$, es **la combinación lineal de las columnas de** A **utilizando las correspondientes entradas en x como pesos**; esto es,

$$A\mathbf{x} = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n\,] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

Observe que $A\mathbf{x}$ está definida sólo si el número de columnas de A es igual al número de entradas en $\mathbf{x}$.

EJEMPLO 1

a. $$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix}$$
$$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

b. $$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$

EJEMPLO 2 Para $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ en $\mathbb{R}^m$, escriba la combinación lineal $3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3$, como una matriz multiplicada por un vector.

Solución Coloque $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ en las columnas de una matriz A, y los pesos 3, −5 y 7 en un vector $\mathbf{x}$. Esto es,

$$3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3 = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} = A\mathbf{x}$$

En la sección 1.3 se aprendió a escribir un sistema de ecuaciones lineales como una ecuación vectorial que implica una combinación lineal de vectores. Por ejemplo, se sabe que el sistema

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 4 \\ -5x_2 + 3x_3 &= 1 \end{aligned} \tag{1}$$

es equivalente a

$$x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ -5 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \tag{2}$$

De igual forma que en el ejemplo 2, la combinación lineal del lado izquierdo puede escribirse como una matriz por un vector, así que (2) se transforma en

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \tag{3}$$

La ecuación (3) tiene la forma $A\mathbf{x} = \mathbf{b}$. Una ecuación como ésta se denomina **ecuación matricial**, para distinguirla de una ecuación vectorial similar a la que se muestra en (2).

Observe cómo la matriz de (3) es simplemente la matriz de coeficientes del sistema (1). Cálculos similares muestran que cualquier sistema de ecuaciones lineales, o cualquier ecuación vectorial del tipo (2), puede escribirse como una ecuación matricial equivalente de la forma $A\mathbf{x} = \mathbf{b}$. Esta simple observación se usará repetidamente a lo largo del texto.

He aquí el resultado formal.

TEOREMA 3

Si A es una matriz de $m \times n$, con columnas $\mathbf{a}_1, \ldots, \mathbf{a}_n$, y si $\mathbf{b}$ está en $\mathbb{R}^m$, la ecuación matricial

$$A\mathbf{x} = \mathbf{b} \tag{4}$$

tiene el mismo conjunto solución que la ecuación vectorial

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{5}$$

la cual, a su vez, tiene el mismo conjunto solución que el sistema de ecuaciones lineales cuya matriz aumentada es

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix} \tag{6}$$

El teorema 3 proporciona una herramienta poderosa para adquirir una mejor percepción de los problemas en álgebra lineal, porque ahora es posible ver un sistema de ecuaciones lineales de tres maneras distintas pero equivalentes: como una ecuación matricial, como una ecuación vectorial o como un sistema de ecuaciones lineales. Cuando se construye un modelo matemático sobre algún problema de la vida real, se tiene la libertad de elegir el punto de vista que resulte más natural. Luego, según sea conveniente, puede cambiarse de una formulación a otra. De cualquier manera, la ecuación matricial, la ecuación vectorial, y el sistema de ecuaciones se resuelven todos de igual forma: reduciendo por filas la matriz aumentada (6). Posteriormente, se analizarán otros métodos de solución.

Existencia de soluciones

La definición de $A\mathbf{x}$ conduce de modo directo al útil enunciado siguiente.

La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución si, y sólo si, $\mathbf{b}$ es una combinación lineal de las columnas de A.

En la sección 1.3, se consideró la pregunta de existencia: "¿Está $\mathbf{b}$ en Gen$\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$?" De manera equivalente: "¿Es consistente $A\mathbf{x} = \mathbf{b}$?" Un problema de existencia aún más difícil es determinar si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente *para toda* $\mathbf{b}$ posible.

EJEMPLO 3 Sea $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. ¿La ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para todas las posibles b_1, b_2, b_3?

Solución Reduzca por filas la matriz aumentada de $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 7 & 5 & b_3 + 3b_1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{bmatrix}$$

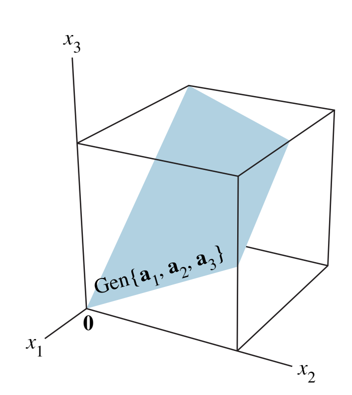


FIGURA 1
Las columnas de $A = [\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{a}_3]$ generan un plano que pasa por $\mathbf{0}$.

La tercera entrada de la columna aumentada es $b_1 - \frac{1}{2}b_2 + b_3$. La ecuación $A\mathbf{x} = \mathbf{b}$ *no* es consistente para toda $\mathbf{b}$ porque algunos valores de $\mathbf{b}$ pueden hacer a $b_1 - \frac{1}{2}b_2 + b_3$ diferente de cero.

La matriz reducida del ejemplo 3 proporciona una descripción de toda $\mathbf{b}$ para la cual la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente: las entradas en $\mathbf{b}$ deben satisfacer

$$b_1 - \tfrac{1}{2}b_2 + b_3 = 0$$

Ésta es la ecuación de un plano que pasa por el origen en $\mathbb{R}^3$. El plano es el conjunto de todas las combinaciones lineales de las tres columnas de A. Vea la figura 1.

La ecuación $A\mathbf{x} = \mathbf{b}$ del ejemplo 3 no cumple con ser consistente para toda $\mathbf{b}$ porque la forma escalonada de A tiene una fila de ceros. Si A tuviera un pivote en cada una de las tres filas, no habría necesidad de hacer los cálculos en la columna aumentada porque, en este caso, una forma escalonada de la matriz aumentada no tendría una fila del tipo $[0 \ \ 0 \ \ 0 \ \ 1]$.

En el teorema siguiente, cuando se afirma que **las columnas de** A **generan** $\mathbb{R}^m$, se pretende establecer que *toda* $\mathbf{b}$ en $\mathbb{R}^m$ es una combinación lineal de las columnas de A. En general, un conjunto de vectores $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^m$ **genera** (o **produce**) $\mathbb{R}^m$ si todo vector en $\mathbb{R}^m$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_p$, esto es, si Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\} = \mathbb{R}^m$.

TEOREMA 4

Sea A una matriz de $m \times n$. Entonces, las siguientes afirmaciones son lógicamente equivalentes. Esto es, para una A en particular, todas estas afirmaciones son verdaderas o todas son falsas.

a. Para cada $\mathbf{b}$ en $\mathbb{R}^m$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución.

b. Cada $\mathbf{b}$ en $\mathbb{R}^m$ es una combinación lineal de las columnas de A.

c. Las columnas de A generan $\mathbb{R}^m$.

d. A tiene una posición pivote en cada fila.

El teorema 4 es uno de los teoremas más útiles de este capítulo. Los enunciados (a), (b) y (c) son equivalentes debido a la definición de $A\mathbf{x}$ y lo que significa para un conjunto de vectores que genera a $\mathbb{R}^m$. El análisis posterior al ejemplo 3 sugiere por qué (a) y (d) son equivalentes; al final de esta sección se demuestra esto. Los ejercicios proporcionarán ejemplos de cómo se utiliza el teorema 4.

Advertencia: El teorema 4 trata acerca de una *matriz de coeficientes*, no de una matriz aumentada. Si una matriz aumentada $[A \quad \mathbf{b}]$ tiene una posición pivote en cada fila, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ puede ser consistente o no.

Cálculo de $A\mathbf{x}$

Los cálculos del ejemplo 1 se basaron en la definición del producto de una matriz A por un vector $\mathbf{x}$. El sencillo ejemplo que se presenta a continuación conducirá a un método más eficiente para calcular las entradas de $A\mathbf{x}$ cuando se resuelvan los problemas a mano.

EJEMPLO 4 Calcule $A\mathbf{x}$. donde $A = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

Solución A partir de la definición,

$$\begin{aligned}
\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= x_1\begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix} \\
&= \begin{bmatrix} 2x_1 \\ -x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix} \qquad (7) \\
&= \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix}
\end{aligned}$$

La primera entrada del producto $A\mathbf{x}$ es una suma de productos (algunas veces llamada *producto punto*), usando la primera fila de A y las entradas de $\mathbf{x}$. Esto es,

$$\begin{bmatrix} 2 & 3 & 4 \\ & & \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ \\ \\ \end{bmatrix}$$

Esta matriz muestra cómo calcular la primera entrada de $A\mathbf{x}$ directamente, sin escribir todos los cálculos que se muestran en (7). De manera similar, la segunda entrada de $A\mathbf{x}$ se puede calcular de inmediato al multiplicar las entradas de la segunda fila de A por las entradas correspondientes de $\mathbf{x}$ para luego sumar los productos resultantes:

$$\begin{bmatrix} & & \\ -1 & 5 & -3 \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \\ -x_1 + 5x_2 - 3x_3 \\ \\ \end{bmatrix}$$

De manera similar, la tercera entrada de $A\mathbf{x}$ se puede calcular a partir de la tercera fila de A y las entradas de $\mathbf{x}$.

REGLA DEL VECTOR FILA PARA CALCULAR $A\mathbf{x}$

Si el producto $A\mathbf{x}$ está definido, entonces la entrada i-ésima de $A\mathbf{x}$ es la suma de los productos de las entradas correspondientes de la fila i de A y del vector $\mathbf{x}$.

EJEMPLO 5

a. $$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1\cdot 4 + 2\cdot 3 + (-1)\cdot 7 \\ 0\cdot 4 + (-5)\cdot 3 + 3\cdot 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

b. $$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 2\cdot 4 + (-3)\cdot 7 \\ 8\cdot 4 + 0\cdot 7 \\ (-5)\cdot 4 + 2\cdot 7 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$

c. $$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} 1\cdot r + 0\cdot s + 0\cdot t \\ 0\cdot r + 1\cdot s + 0\cdot t \\ 0\cdot r + 0\cdot s + 1\cdot t \end{bmatrix} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}$$

Por definición, la matriz del ejemplo 5(c) con números 1 en la diagonal y ceros en los demás lugares se denomina **matriz identidad**, y se denota mediante I. Los cálculos del inciso (c) muestran que $I\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^3$. Existe una matriz identidad análoga de $n \times n$ que algunas veces se escribe como I_n. Al igual que en el inciso (c), $I_n\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ presente en $\mathbb{R}^n$.

Propiedades del producto matriz-vector $A\mathbf{x}$

Los enunciados del teorema siguiente son importantes y se usarán a lo largo del texto. La demostración está basada en la definición de $A\mathbf{x}$ y en las propiedades algebraicas de $\mathbb{R}^n$.

TEOREMA 5

Si A es una matriz de $m \times n$, $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$, y c es un escalar, entonces

a. $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;
b. $A(c\mathbf{u}) = c(A\mathbf{u})$

DEMOSTRACIÓN En aras de la simplicidad, tome $n = 3$, $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$, y $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^3$. (La demostración del caso general es similar.) Para $i = 1, 2, 3$, sean u_i y v_i las i-ésimas entradas de $\mathbf{u}$ y $\mathbf{v}$, respectivamente. Para demostrar el enunciado (a), calcule $A(\mathbf{u} + \mathbf{v})$ como una combinación lineal de las columnas de A usando como pesos las entradas de $\mathbf{u} + \mathbf{v}$.

$$A(\mathbf{u}+\mathbf{v}) = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{bmatrix}$$

Entradas en $\mathbf{u}+\mathbf{v}$

$$= (u_1+v_1)\mathbf{a}_1 + (u_2+v_2)\mathbf{a}_2 + (u_3+v_3)\mathbf{a}_3$$

Columnas de A

$$= (u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) + (v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3)$$
$$= A\mathbf{u} + A\mathbf{v}$$

Para demostrar el enunciado (b), calcule $A(c\mathbf{u})$ como una combinación lineal de las columnas de A usando como pesos las entradas de $c\mathbf{u}$.

$$A(c\mathbf{u}) = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]\begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} = (cu_1)\mathbf{a}_1 + (cu_2)\mathbf{a}_2 + (cu_3)\mathbf{a}_3$$
$$= c(u_1\mathbf{a}_1) + c(u_2\mathbf{a}_2) + c(u_3\mathbf{a}_3)$$
$$= c(u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3)$$
$$= c(A\mathbf{u})$$

■

NOTA NUMÉRICA

Si se desea optimizar un algoritmo de computadora para calcular $A\mathbf{x}$, la secuencia de cálculos debe incluir datos almacenados en posiciones contiguas de memoria. Los algoritmos profesionales que más se usan para cálculos de matrices están escritos en Fortran, un lenguaje que almacena una matriz como un conjunto de columnas. Tales algoritmos calculan $A\mathbf{x}$ como una combinación lineal de las columnas de A. En contraste, si un programa está escrito en el popular lenguaje C, que almacena las matrices por filas, $A\mathbf{x}$ deberá calcularse mediante la regla alternativa que utiliza las filas de A.

DEMOSTRACIÓN DEL TEOREMA 4 Tal como se definió después del teorema 4, los enunciados (a), (b) y (c) son lógicamente equivalentes. Entonces, resulta suficiente demostrar (para una matriz A arbitraria) que (a) y (d) son ambos verdaderos o ambos falsos. En tal caso, los cuatro enunciados serán todos ciertos o todos falsos.

Sea U una forma escalonada de A. Dado $\mathbf{b}$ en $\mathbb{R}^m$, la matriz aumentada $[A \quad \mathbf{b}]$ se puede reducir por filas a una matriz aumentada $[U \quad \mathbf{d}]$ para alguna $\mathbf{d}$ presente en $\mathbb{R}^m$:

$$[\,A \quad \mathbf{b}\,] \sim \cdots \sim [\,U \quad \mathbf{d}\,]$$

Si el enunciado (d) es cierto, entonces cada fila de U contiene una posición pivote y no puede haber pivotes en la columna aumentada. Así que $A\mathbf{x} = \mathbf{b}$ tiene una solución para cualquier $\mathbf{b}$ y (a) es verdadero. Si (d) es falso, la última fila de U es de sólo ceros. Sea $\mathbf{d}$ cualquier vector con un 1 en su última entrada. Entonces $[U \quad \mathbf{d}]$ representa un sistema *inconsistente*. Como las operaciones por filas son reversibles, $[U \quad \mathbf{d}]$ puede transformarse a la forma $[A \quad \mathbf{b}]$. El nuevo sistema $A\mathbf{x} = \mathbf{b}$ también es inconsistente, y (a) es falso. ■

PROBLEMAS DE PRÁCTICA

1. Sea $A = \begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}$, y $\mathbf{b} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$. Se puede mostrar que $\mathbf{p}$ es una solución de $A\mathbf{x} = \mathbf{b}$. Utilice este hecho para mostrar a $\mathbf{b}$ como una combinación lineal específica de las columnas de A.

2. Sean $A = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$, y $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. Verifique el teorema 5(a) calculando, en este caso, $A(\mathbf{u} + \mathbf{v})$ y $A\mathbf{u} + A\mathbf{v}$.

1.4 EJERCICIOS

Encuentre los productos en los ejercicios 1 a 4 usando (a) la definición, como en el ejemplo 1, y (b) la regla del vector fila para calcular $A\mathbf{x}$. Si un producto no está definido, explique por qué.

1. $\begin{bmatrix} -4 & 2 \\ 1 & 6 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ 7 \end{bmatrix}$

2. $\begin{bmatrix} 2 \\ 6 \\ -1 \end{bmatrix}\begin{bmatrix} 5 \\ -1 \end{bmatrix}$

3. $\begin{bmatrix} 6 & 5 \\ -4 & -3 \\ 7 & 6 \end{bmatrix}\begin{bmatrix} 2 \\ -3 \end{bmatrix}$

4. $\begin{bmatrix} 8 & 3 & -4 \\ 5 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 5 a 8, use la definición de $A\mathbf{x}$ para escribir la ecuación matricial como una ecuación vectorial, o viceversa.

5. $\begin{bmatrix} 5 & 1 & -8 & 4 \\ -2 & -7 & 3 & -5 \end{bmatrix}\begin{bmatrix} 5 \\ -1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -8 \\ 16 \end{bmatrix}$

6. $\begin{bmatrix} 7 & -3 \\ 2 & 1 \\ 9 & -6 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} -2 \\ -5 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ 12 \\ -4 \end{bmatrix}$

7. $x_1\begin{bmatrix} 4 \\ -1 \\ 7 \\ -4 \end{bmatrix} + x_2\begin{bmatrix} -5 \\ 3 \\ -5 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} 7 \\ -8 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ -8 \\ 0 \\ -7 \end{bmatrix}$

8. $z_1\begin{bmatrix} 4 \\ -2 \end{bmatrix} + z_2\begin{bmatrix} -4 \\ 5 \end{bmatrix} + z_3\begin{bmatrix} -5 \\ 4 \end{bmatrix} + z_4\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 13 \end{bmatrix}$

En los ejercicios 9 y 10, primero escriba el sistema como una ecuación vectorial y después como una ecuación matricial.

9. $\begin{aligned} 3x_1 + x_2 - 5x_3 &= 9 \\ x_2 + 4x_3 &= 0 \end{aligned}$

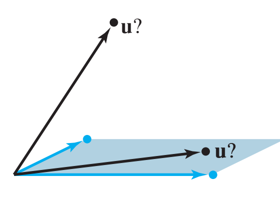
10. $\begin{aligned} 8x_1 - x_2 &= 4 \\ 5x_1 + 4x_2 &= 1 \\ x_1 - 3x_2 &= 2 \end{aligned}$

Dados A y $\mathbf{b}$ en los ejercicios 11 y 12, escriba la matriz aumentada para el sistema lineal que corresponde a la ecuación matricial $A\mathbf{x} = \mathbf{b}$. Después resuelva el sistema y escriba la solución como un vector.

11. $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 5 \\ -2 & -4 & -3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ 2 \\ 9 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 2 & 1 \\ -3 & -1 & 2 \\ 0 & 5 & 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$

13. Sean $\mathbf{u} = \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix}$ y $A = \begin{bmatrix} 3 & -5 \\ -2 & 6 \\ 1 & 1 \end{bmatrix}$. ¿Está $\mathbf{u}$ en el plano en $\mathbb{R}^3$ generado por las columnas de A? (Vea la figura.) ¿Por qué sí o por qué no?



¿Dónde está $\mathbf{u}$?

14. Sean $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 2 \end{bmatrix}$ y $A = \begin{bmatrix} 5 & 8 & 7 \\ 0 & 1 & -1 \\ 1 & 3 & 0 \end{bmatrix}$. ¿Está $\mathbf{u}$ en el subconjunto en $\mathbb{R}^3$ generado por las columnas de A? ¿Por qué sí o por qué no?

15. Sean $A = \begin{bmatrix} 2 & -1 \\ -6 & 3 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Muestre que la ecuación $A\mathbf{x} = \mathbf{b}$ no tiene una solución para todas las $\mathbf{b}$ posibles, y describa el conjunto de todas las $\mathbf{b}$ para las cuales $A\mathbf{x} = \mathbf{b}$ *sí* tiene una solución.

16. Repita el ejercicio 15: $A = \begin{bmatrix} 1 & -3 & -4 \\ -3 & 2 & 6 \\ 5 & -1 & -8 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

Los ejercicios 17 a 20 se refieren a las matrices A y B que se presentan a continuación. Realice los cálculos adecuados que justifiquen sus respuestas y mencione un teorema apropiado.

$$A = \begin{bmatrix} 1 & 3 & 0 & 3 \\ -1 & -1 & -1 & 1 \\ 0 & -4 & 2 & -8 \\ 2 & 0 & 3 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 3 & -2 & 2 \\ 0 & 1 & 1 & -5 \\ 1 & 2 & -3 & 7 \\ -2 & -8 & 2 & -1 \end{bmatrix}$$

17. ¿Cuántas filas de A contienen una posición pivote? ¿La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^4$?

18. ¿Las columnas de B generan $\mathbb{R}^4$? ¿La ecuación $B\mathbf{x} = \mathbf{y}$ tiene una solución para cada $\mathbf{y}$ en $\mathbb{R}^4$?

19. ¿Puede escribirse cada vector en $\mathbb{R}^4$ como una combinación lineal de las columnas de la matriz A? ¿Las columnas de A generan $\mathbb{R}^4$?

20. ¿Puede escribirse cada vector en $\mathbb{R}^4$ como una combinación lineal de las columnas de la matriz B? ¿Las columnas de B generan $\mathbb{R}^3$?

21. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$.

¿Genera $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a $\mathbb{R}^4$? ¿Por qué sí o por qué no?

22. Sean $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -3 \\ 8 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ -1 \\ -5 \end{bmatrix}$.

¿Genera $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a $\mathbb{R}^3$? ¿Por qué sí o por qué no?

En los ejercicios 23 y 24, señale cada enunciado como verdadero o falso. Justifique cada respuesta.

23. a. La ecuación $A\mathbf{x} = \mathbf{b}$ se conoce como una *ecuación vectorial*.

b. Un vector $\mathbf{b}$ es una combinación lineal de las columnas de una matriz A si, y sólo si, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución.

c. La ecuación $A\mathbf{x} = \mathbf{b}$ es consistente si la matriz aumentada $[A \ \ \mathbf{b}]$ tiene una posición pivote en cada fila.

d. La primera entrada en el producto $A\mathbf{x}$ es una suma de productos.

e. Si las columnas de una matriz A de $m \times n$ generan $\mathbb{R}^m$, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para cada $\mathbf{b}$ presente en $\mathbb{R}^m$.

f. Si A es una matriz de $m \times n$ y la ecuación $A\mathbf{x} = \mathbf{b}$ es inconsistente para alguna $\mathbf{b}$ en $\mathbb{R}^m$, entonces A no puede tener una posición pivote en cada fila.

24. a. Cualquier ecuación matricial $A\mathbf{x} = \mathbf{b}$ corresponde a una ecuación vectorial con el mismo conjunto solución.

b. Cualquier combinación lineal de vectores siempre puede escribirse en la forma $A\mathbf{x}$ para una matriz A y un vector $\mathbf{x}$.

c. El conjunto solución de un sistema lineal cuya matriz aumentada es $[\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{a}_3 \ \ \mathbf{b}]$ es el mismo que el conjunto solución de $A\mathbf{x} = \mathbf{b}$, si $A = [\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{a}_3]$

d. Si la ecuación $A\mathbf{x} = \mathbf{b}$ es inconsistente, entonces $\mathbf{b}$ no está en el conjunto generado por las columnas de A.

e. Si la matriz aumentada $[A \ \ \mathbf{b}]$ tiene una posición pivote en cada fila, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es inconsistente.

f. Si A es una matriz $m \times n$ cuyas columnas no generan $\mathbb{R}^m$, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es inconsistente para alguna $\mathbf{b}$ en $\mathbb{R}^m$.

25. Observe que $\begin{bmatrix} 4 & -3 & 1 \\ 5 & -2 & 5 \\ -6 & 2 & -3 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -7 \\ -3 \\ 10 \end{bmatrix}$. Use este hecho (y no realice operaciones de fila) para encontrar los escalares c_1, c_2, c_3 tales que

$$\begin{bmatrix} -7 \\ -3 \\ 10 \end{bmatrix} = c_1 \begin{bmatrix} 4 \\ 5 \\ -6 \end{bmatrix} + c_2 \begin{bmatrix} -3 \\ -2 \\ 2 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}.$$

26. Sean $\mathbf{u} = \begin{bmatrix} 7 \\ 2 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$, y $\mathbf{w} = \begin{bmatrix} 6 \\ 1 \\ 0 \end{bmatrix}$.

Se puede demostrar que $3\mathbf{u} - 5\mathbf{v} - \mathbf{w} = \mathbf{0}$. Utilice este hecho (y no realice operaciones de fila) para encontrar las x_1 y x_2 que satisfagan la ecuación

$$\begin{bmatrix} 7 & 3 \\ 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ 0 \end{bmatrix}.$$

27. Sean $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ y $\mathbf{v}$ vectores en $\mathbb{R}^5$, y sean x_1, x_2 y x_3 escalares. Escriba la siguiente ecuación vectorial como una ecuación matricial. Identifique cualquier símbolo que decida utilizar.

$$x_1\mathbf{q}_1 + x_2\mathbf{q}_2 + x_3\mathbf{q}_3 = \mathbf{v}$$

28. Reescriba la siguiente ecuación matricial (numérica) en forma simbólica como una ecuación vectorial, y utilice los símbolos $\mathbf{v}_1, \mathbf{v}_2, \ldots$ para los vectores y $c_1, c_2, \ldots$ para los escalares.

Defina lo que representa cada símbolo usando los datos de la ecuación matricial.

$$\begin{bmatrix} -3 & 5 & -4 & 9 & 7 \\ 5 & 8 & 1 & -2 & -4 \end{bmatrix} \begin{bmatrix} -3 \\ 2 \\ 4 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ -1 \end{bmatrix}$$

29. Construya una matriz de 3 × 3, en forma no escalonada, cuyas columnas generen $\mathbb{R}^3$. Muestre que dicha matriz tiene la propiedad deseada.

30. Construya una matriz 3 × 3, en forma no escalonada, cuyas columnas *no* generen $\mathbb{R}^3$. Muestre que dicha matriz tiene la propiedad deseada.

31. Sea A una matriz de 3 × 2. Explique por qué la ecuación $A\mathbf{x} = \mathbf{b}$ no puede ser consistente para toda $\mathbf{b}$ en $\mathbb{R}^3$. Generalice su argumento para el caso de una A arbitraria con más filas que columnas.

32. ¿Podría un conjunto de tres vectores en $\mathbb{R}^4$ generar todo $\mathbb{R}^4$? Explique su respuesta. ¿Qué sucede con n vectores en $\mathbb{R}^m$ cuando n es menor que m?

33. Suponga que A es una matriz de 4×3 y $\mathbf{b}$ un vector en $\mathbb{R}^4$ con la propiedad de que $A\mathbf{x} = \mathbf{b}$ tiene una solución única. ¿Qué puede decirse acerca de la forma escalonada reducida de A? Justifique su respuesta.

34. Suponga que A es una matriz de 3×3 y $\mathbf{b}$ un vector en $\mathbb{R}^3$ con la propiedad de que $A\mathbf{x} = \mathbf{b}$ tiene una solución única. Explique por qué las columnas de A deben generar a $\mathbb{R}^3$.

35. Sean A una matriz de 3×4, $\mathbf{y}_1$ e $\mathbf{y}_2$ vectores en $\mathbb{R}^3$, y $\mathbf{w} = \mathbf{y}_1 + \mathbf{y}_2$. Suponga que $\mathbf{y}_1 = A\mathbf{x}_1$ y $\mathbf{y}_2 = A\mathbf{x}_2$ para algunos vectores $\mathbf{x}_1$ y $\mathbf{x}_2$ en $\mathbb{R}^4$. ¿Cuál hecho permite concluir que el sistema $A\mathbf{x} = \mathbf{w}$ es consistente? (*Nota:* $\mathbf{x}_1$ y $\mathbf{x}_2$ denotan vectores, no entradas escalares de vectores.)

36. Sean A una matriz de 5×3, $\mathbf{y}$ un vector en $\mathbb{R}^3$, y $\mathbf{z}$ un vector en $\mathbb{R}^5$. Suponga que $A\mathbf{y} = \mathbf{z}$. ¿Cuál hecho permite concluir que el sistema $A\mathbf{x} = 4\mathbf{z}$ es consistente?

[M] En los ejercicios 37 a 40, determine si las columnas de la matriz generan a $\mathbb{R}^4$.

37. $\begin{bmatrix} 7 & 2 & -5 & 8 \\ -5 & -3 & 4 & -9 \\ 6 & 10 & -2 & 7 \\ -7 & 9 & 2 & 15 \end{bmatrix}$

38. $\begin{bmatrix} 5 & -7 & -4 & 9 \\ 6 & -8 & -7 & 5 \\ 4 & -4 & -9 & -9 \\ -9 & 11 & 16 & 7 \end{bmatrix}$

39. $\begin{bmatrix} 12 & -7 & 11 & -9 & 5 \\ -9 & 4 & -8 & 7 & -3 \\ -6 & 11 & -7 & 3 & -9 \\ 4 & -6 & 10 & -5 & 12 \end{bmatrix}$

40. $\begin{bmatrix} 8 & 11 & -6 & -7 & 13 \\ -7 & -8 & 5 & 6 & -9 \\ 11 & 7 & -7 & -9 & -6 \\ -3 & 4 & 1 & 8 & 7 \end{bmatrix}$

41. **[M]** En la matriz del ejercicio 39, encuentre una columna que se pueda borrar sin que las columnas restantes dejen de generar a $\mathbb{R}^4$.

42. **[M]** En la matriz del ejercicio 40, encuentre una columna que se pueda borrar sin que las columnas restantes dejen de generar a $\mathbb{R}^4$. ¿Podría borrarse más de una columna?

SG Dominio de los conceptos de álgebra lineal: secciones 1 a 19 (Mastering Linear Algebra Concepts: Span 1-19)

CD Resolución de $A\mathbf{x} = \mathbf{b}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. La ecuación matricial

$$\begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$$

es equivalente a la ecuación vectorial

$$3\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} - 2\begin{bmatrix} 5 \\ 1 \\ -8 \end{bmatrix} + 0\begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix} - 4\begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$$

que expresa $\mathbf{b}$ como una combinación lineal de las columnas de A.

2.
$$\mathbf{u}+\mathbf{v}=\begin{bmatrix}4\\-1\end{bmatrix}+\begin{bmatrix}-3\\5\end{bmatrix}=\begin{bmatrix}1\\4\end{bmatrix}$$
$$A(\mathbf{u}+\mathbf{v})=\begin{bmatrix}2&5\\3&1\end{bmatrix}\begin{bmatrix}1\\4\end{bmatrix}=\begin{bmatrix}2+20\\3+4\end{bmatrix}=\begin{bmatrix}22\\7\end{bmatrix}$$
$$A\mathbf{u}+A\mathbf{v}=\begin{bmatrix}2&5\\3&1\end{bmatrix}\begin{bmatrix}4\\-1\end{bmatrix}+\begin{bmatrix}2&5\\3&1\end{bmatrix}\begin{bmatrix}-3\\5\end{bmatrix}$$
$$=\begin{bmatrix}3\\11\end{bmatrix}+\begin{bmatrix}19\\-4\end{bmatrix}=\begin{bmatrix}22\\7\end{bmatrix}$$

1.5 CONJUNTOS SOLUCIÓN DE LOS SISTEMAS LINEALES

En álgebra lineal, los conjuntos solución de los sistemas lineales son objetos de estudio importantes. Posteriormente aparecerán en diversos contextos. Esta sección utiliza notación vectorial para dar descripciones explícitas y geométricas de dichos conjuntos solución.

Sistemas lineales homogéneos

Se dice que un sistema de ecuaciones lineales es **homogéneo** si se puede escribir en la forma $A\mathbf{x}=\mathbf{0}$, donde A es una matriz de $m\times n$ y $\mathbf{0}$ es el vector cero en $\mathbb{R}^m$. Un sistema $A\mathbf{x}=\mathbf{0}$ como éste *siempre* tiene al menos una solución, a saber, $\mathbf{x}=\mathbf{0}$ (el vector cero en $\mathbb{R}^n$). Por lo general, esta solución cero se denomina **solución trivial**. Para una ecuación dada $A\mathbf{x}=\mathbf{0}$, la pregunta importante es si existe o no una **solución no trivial**, esto es, un vector $\mathbf{x}$ diferente de cero que satisfaga $A\mathbf{x}=\mathbf{0}$. El teorema de existencia y unicidad de la sección 1.2 (teorema 2) conduce de inmediato al siguiente enunciado.

> La ecuación homogénea $A\mathbf{x}=\mathbf{0}$ tiene una solución no trivial si, y sólo si, la ecuación tiene por lo menos una variable libre.

EJEMPLO 1 Determine si el siguiente sistema homogéneo tiene una solución no trivial. Después describa el conjunto solución.

$$\begin{aligned}3x_1+5x_2-4x_3&=0\\-3x_1-2x_2+4x_3&=0\\6x_1+x_2-8x_3&=0\end{aligned}$$

Solución Sea A la matriz de coeficientes del sistema, reduzca por filas la matriz aumentada $[A\quad\mathbf{0}]$ a la forma escalonada.

$$\begin{bmatrix}3&5&-4&0\\-3&-2&4&0\\6&1&-8&0\end{bmatrix}\sim\begin{bmatrix}3&5&-4&0\\0&3&0&0\\0&-9&0&0\end{bmatrix}\sim\begin{bmatrix}3&5&-4&0\\0&3&0&0\\0&0&0&0\end{bmatrix}$$

Como x_3 es una variable libre, $A\mathbf{x} = \mathbf{0}$ tiene soluciones no triviales (una para cada valor de x_3). Para describir el conjunto solución, continúe la reducción por filas de $[A \quad \mathbf{0}]$ hasta la forma escalonada *reducida*:

$$\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_1 \quad - \tfrac{4}{3}x_3 &= 0 \\ x_2 \qquad &= 0 \\ 0 &= 0 \end{aligned}$$

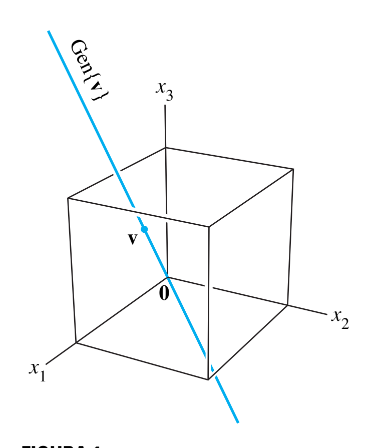


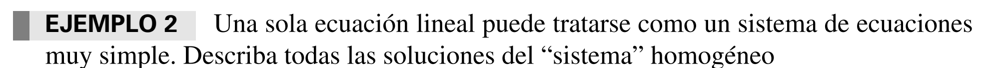
FIGURA 1

Resuelva para las variables básicas x_1 y x_2 y obtenga $x_1 = \frac{4}{3}x_3$, $x_2 = 0$, con x_3 libre. Como vector, la solución general de $A\mathbf{x} = \mathbf{0}$ tiene la forma

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} = x_3\mathbf{v}, \quad \text{donde } \mathbf{v} = \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

Aquí x_3 se factoriza para obtener el vector solución general. Esto muestra que cada solución de $A\mathbf{x} = \mathbf{0}$ en este caso es un múltiplo escalar de $\mathbf{v}$. La solución trivial se obtiene al seleccionar $x_3 = 0$. Geométricamente, el conjunto solución es una línea en $\mathbb{R}^3$ que pasa por $\mathbf{0}$. Vea la figura 1.

Observe que una solución no trivial $\mathbf{x}$ puede tener algunas entradas iguales a cero, pero no todas.



EJEMPLO 2 Una sola ecuación lineal puede tratarse como un sistema de ecuaciones muy simple. Describa todas las soluciones del "sistema" homogéneo

$$10x_1 - 3x_2 - 2x_3 = 0 \tag{1}$$

Solución No hay necesidad de emplear notación matricial. Resuelva para la variable básica x_1 en términos de las variables libres. La solución general es $x_1 = .3x_2 + .2x_3$ con x_2 y x_3 libres. Como vector, la solución general es

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 + .2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} .2x_3 \\ 0 \\ x_3 \end{bmatrix}$$

$$= x_2 \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} .3 \\ 1 \\ 0 \end{bmatrix}} + x_3 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} .2 \\ 0 \\ 1 \end{bmatrix}} \quad (\text{con } x_2, x_3 \text{ libres}) \tag{2}$$

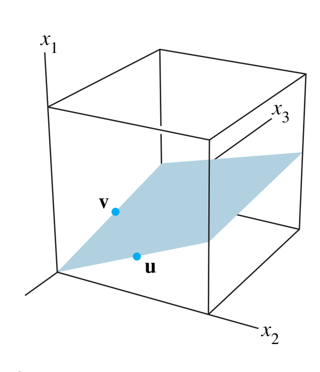


FIGURA 2

Este cálculo muestra que cada solución de (1) es una combinación lineal de los vectores $\mathbf{u}$ y $\mathbf{v}$, los cuales se muestran en (2). Esto es, el conjunto solución es Gen$\{\mathbf{u}, \mathbf{v}\}$. Ya que ni $\mathbf{u}$ ni $\mathbf{v}$ son múltiplos escalares entre sí, el conjunto solución es un plano que pasa por el origen. Vea la figura 2.

Los ejemplos 1 y 2, junto con los ejercicios, ilustran el hecho de que el conjunto solución de una ecuación homogénea $A\mathbf{x} = \mathbf{0}$ siempre puede expresarse explícitamente como Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ para vectores adecuados $\mathbf{v}_1, \ldots, \mathbf{v}_p$. Si la única solución es el

vector cero, entonces el conjunto solución es Gen{**0**}. Si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene sólo una variable libre, el conjunto solución es una línea que pasa por el origen, como en la figura 1. Un plano que pasa por el origen, como en la figura 2, proporciona una buena imagen mental del conjunto solución de $A\mathbf{x} = \mathbf{0}$ cuando hay dos o más variables libres. Sin embargo, debe observarse que es posible usar una figura similar para visualizar Gen{**u**, **v**} aunque **u** y **v** no surjan como soluciones de $A\mathbf{x} = \mathbf{0}$. Vea la figura 11 en la sección 1.3.

Forma vectorial paramétrica

La ecuación original (1) para el plano del ejemplo 2 es una descripción *implícita* del plano. Resolver esta ecuación equivale a encontrar una descripción *explícita* del plano como el conjunto generado por **u** y **v**. La ecuación (2) se llama **ecuación vectorial paramétrica** del plano. Una ecuación de este tipo se escribe algunas veces como

$$\mathbf{x} = s\mathbf{u} + t\mathbf{v} \quad (s, t \text{ en } \mathbb{R})$$

para enfatizar que los parámetros varían sobre todos los números reales. En el ejemplo 1 la ecuación $\mathbf{x} = x_3\mathbf{v}$ (con x_3 libre), o $\mathbf{x} = t\mathbf{v}$ (con t en $\mathbb{R}$), es una ecuación vectorial paramétrica de una recta. Siempre que un conjunto solución se describa explícitamente con vectores, como en los ejemplos 1 y 2, se dirá que la solución está en **forma vectorial paramétrica**.

Soluciones de sistemas no homogéneos

Cuando un sistema lineal no homogéneo tiene muchas soluciones, la solución general puede escribirse en forma vectorial paramétrica como un vector más una combinación lineal arbitraria de vectores que satisfaga el sistema homogéneo correspondiente.

EJEMPLO 3 Describa todas las soluciones de $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}$$

Solución Aquí, A es la matriz de coeficientes del ejemplo 1. Las operaciones de fila en $[A \quad \mathbf{b}]$ producen

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 \quad\quad - \tfrac{4}{3}x_3 &= -1 \\ x_2 \quad\quad &= 2 \\ 0 &= 0 \end{aligned}$$

Así que $x_1 = -1 + \frac{4}{3}x_3$, $x_2 = 2$, y x_3 es libre. Como vector, la solución general de $A\mathbf{x} = \mathbf{b}$ tiene la forma

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 + \frac{4}{3}x_3 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}}_{\mathbf{p}} + x_3 \underbrace{\begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}}$$

La ecuación $\mathbf{x} = \mathbf{p} + x_3\mathbf{v}$, o bien, escribiendo t como un parámetro general,

$$\mathbf{x} = \mathbf{p} + t\mathbf{v} \quad (t \text{ en } \mathbb{R}) \tag{3}$$

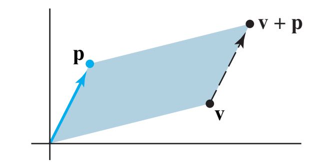


FIGURA 3
La suma de **p** más **v** traslada **v** a **v** + **p**.

describe el conjunto solución de $A\mathbf{x} = \mathbf{b}$ en forma vectorial paramétrica. Del ejemplo 1, recuerde que el conjunto solución de $A\mathbf{x} = \mathbf{0}$ tiene la ecuación vectorial paramétrica

$$\mathbf{x} = t\mathbf{v} \quad (t \text{ en } \mathbb{R}) \tag{4}$$

[con la misma **v** que aparece en (3)]. Así, las soluciones de $A\mathbf{x} = \mathbf{b}$ se obtienen sumando el vector **p** a las soluciones de $A\mathbf{x} = \mathbf{0}$. El mismo vector **p** es sólo una solución particular de $A\mathbf{x} = \mathbf{b}$ [que corresponde a $t = 0$ en (3)].

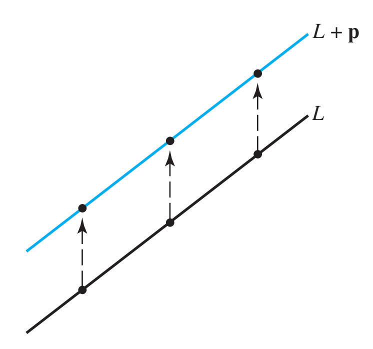


FIGURA 4
Línea trasladada.

Para describir geométricamente el conjunto solución de $A\mathbf{x} = \mathbf{b}$, la suma vectorial puede verse como una *traslación.* Dados **v** y **p** en $\mathbb{R}^2$ o $\mathbb{R}^3$, el efecto de sumar **p** a **v** es *mover* a **v** en una dirección que sea paralela a la línea que pasa por **p** y **0**. Se dice que **p** traslada **a v** hasta **v** + **p**. Vea la figura 3. Si cada punto sobre una línea L en $\mathbb{R}^2$ o $\mathbb{R}^3$ se traslada mediante un vector **p**, el resultado es una línea paralela a L. Vea la figura 4.

Suponga que L es la línea que pasa por **0** y **v**, descrita mediante la ecuación (4). Al sumar **p** a cada punto de L se produce la línea trasladada descrita con la ecuación (3). Observe que **p** está sobre la línea (3). A (3) se le conoce como **la ecuación de la línea que pasa por p paralela a v**. Así, *el conjunto solución de* $A\mathbf{x} = \mathbf{b}$ *es una línea que pasa por* **p** *y es paralela al conjunto solución de* $A\mathbf{x} = \mathbf{0}$. En la figura 5 se ilustra este caso.

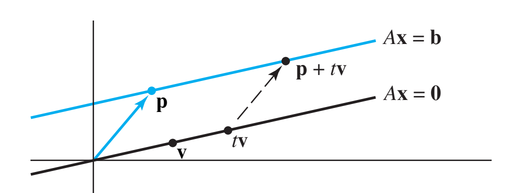


FIGURA 5 Conjuntos solución paralelos de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$.

La relación entre los conjuntos solución de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$ mostrada en la figura 5 se generaliza para cualquier ecuación *consistente* $A\mathbf{x} = \mathbf{b}$, aunque el conjunto solución será más grande que una línea cuando haya varias variables libres. El teorema siguiente proporciona el enunciado preciso. En el ejercicio 25 se da una demostración.

TEOREMA 6

Suponga que la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para alguna **b** dada, y sea **p** una solución. Entonces el conjunto solución de $A\mathbf{x} = \mathbf{b}$ es el conjunto de todos los vectores de la forma $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, donde $\mathbf{v}_h$ es cualquier solución de la ecuación homogénea $A\mathbf{x} = \mathbf{0}$.

El teorema 6 establece que si $A\mathbf{x} = \mathbf{b}$ tiene una solución, entonces el conjunto solución se obtiene al trasladar el conjunto solución de $A\mathbf{x} = \mathbf{0}$; usando cualquier solución particular **p** de $A\mathbf{x} = \mathbf{b}$ para efectuar la traslación. La figura 6 ilustra el caso en que existen dos variables libres. Aun cuando $n > 3$, la imagen mental del conjunto solución

de un sistema consistente $A\mathbf{x} = \mathbf{b}$ (con $\mathbf{b} \neq 0$) es un solo punto diferente de cero, o bien una línea o un plano que no pasa por el origen.

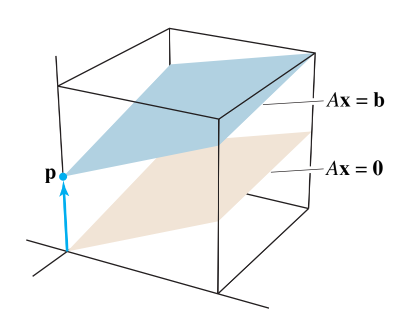


FIGURA 6 Conjuntos solución paralelos de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$.

Advertencia: El teorema 6 y la figura 6 se aplican solamente a una ecuación $A\mathbf{x} = \mathbf{b}$ que tenga por lo menos una solución $\mathbf{p}$ diferente de cero. Cuando $A\mathbf{x} = \mathbf{b}$ no tiene solución, el conjunto solución está vacío.

El algoritmo siguiente describe en términos generales los cálculos mostrados en los ejemplos 1, 2 y 3.

ESCRITURA DE UN CONJUNTO SOLUCIÓN (DE UN SISTEMA CONSISTENTE) EN FORMA VECTORIAL PARAMÉTRICA

1. Reduzca por filas la matriz aumentada a la forma escalonada reducida.
2. Exprese cada variable básica en términos de cualesquiera variables libres que aparezcan en una ecuación.
3. Escriba una solución típica $\mathbf{x}$ como un vector cuyas entradas dependan de las variables libres, si éstas existen.
4. Descomponga $\mathbf{x}$ en una combinación lineal de vectores (con entradas numéricas) usando como parámetros las variables libres.

PROBLEMAS DE PRÁCTICA

1. Cada una de las siguientes ecuaciones determina un plano en $\mathbb{R}^3$. ¿Se intersecan los dos planos? Si lo hacen, describa su intersección.

$$\begin{aligned} x_1 + 4x_2 - 5x_3 &= 0 \\ 2x_1 - x_2 + 8x_3 &= 9 \end{aligned}$$

2. Escriba la solución general de $10x_1 - 3x_2 - 2x_3 = 7$ en forma vectorial paramétrica, y relacione el conjunto solución con el que se encontró en el ejemplo 2.

1.5 EJERCICIOS

En los ejercicios 1 a 4, determine si el sistema tiene una solución no trivial. Trate de emplear tan pocas operaciones por fila como sea posible.

1. $$\begin{aligned} 2x_1 - 5x_2 + 8x_3 &= 0 \\ -2x_1 - 7x_2 + x_3 &= 0 \\ 4x_1 + 2x_2 + 7x_3 &= 0 \end{aligned}$$

2. $$\begin{aligned} x_1 - 3x_2 + 7x_3 &= 0 \\ -2x_1 + x_2 - 4x_3 &= 0 \\ x_1 + 2x_2 + 9x_3 &= 0 \end{aligned}$$

3. $$\begin{aligned} -3x_1 + 5x_2 - 7x_3 &= 0 \\ -6x_1 + 7x_2 + x_3 &= 0 \end{aligned}$$

4. $$\begin{aligned} -5x_1 + 7x_2 + 9x_3 &= 0 \\ x_1 - 2x_2 + 6x_3 &= 0 \end{aligned}$$

En los ejercicios 5 y 6, siga el método de los ejemplos 1 y 2 para escribir el conjunto solución del sistema homogéneo dado en su forma vectorial paramétrica.

5. $$\begin{aligned} x_1 + 3x_2 + x_3 &= 0 \\ -4x_1 - 9x_2 + 2x_3 &= 0 \\ - 3x_2 - 6x_3 &= 0 \end{aligned}$$

6. $$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 0 \\ x_1 + 4x_2 - 8x_3 &= 0 \\ -3x_1 - 7x_2 + 9x_3 &= 0 \end{aligned}$$

En los ejercicios 7 a 12, describa todas las soluciones de $A\mathbf{x} = \mathbf{0}$ en forma vectorial paramétrica, donde A sea equivalente por filas a la matriz dada.

7. $\begin{bmatrix} 1 & 3 & -3 & 7 \\ 0 & 1 & -4 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 1 & -2 & -9 & 5 \\ 0 & 1 & 2 & -6 \end{bmatrix}$

9. $\begin{bmatrix} 3 & -9 & 6 \\ -1 & 3 & -2 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 3 & 0 & -4 \\ 2 & 6 & 0 & -8 \end{bmatrix}$

11. $\begin{bmatrix} 1 & -4 & -2 & 0 & 3 & -5 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 5 & 2 & -6 & 9 & 0 \\ 0 & 0 & 1 & -7 & 4 & -8 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

13. Suponga que el conjunto solución de cierto sistema de ecuaciones lineales puede describirse como $x_1 = 5 + 4x_3$, $x_2 = -2 - 7x_3$, con x_3 libre. Use vectores para describir este conjunto como una línea en $\mathbb{R}^3$.

14. Suponga que el conjunto solución de cierto sistema de ecuaciones lineales puede describirse como $x_1 = 3x_4$, $x_2 = 8 + x_4$, $x_3 = 2 - 5x_4$, con x_4 libre. Use vectores para describir este conjunto como una "línea" en $\mathbb{R}^4$.

15. Siga el método del ejemplo 3 para describir las soluciones del siguiente sistema. También, proporcione una descripción geométrica del conjunto solución y compárelo con el del ejercicio 5.

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 1 \\ -4x_1 - 9x_2 + 2x_3 &= -1 \\ - 3x_2 - 6x_3 &= -3 \end{aligned}$$

16. Igual que en el ejercicio 15, describa las soluciones del siguiente sistema en forma vectorial paramétrica y proporcione una comparación geométrica con el conjunto solución del ejercicio 6.

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 4 \\ x_1 + 4x_2 - 8x_3 &= 7 \\ -3x_1 - 7x_2 + 9x_3 &= -6 \end{aligned}$$

17. Describa y compare los conjuntos solución *de* $x_1 + 9x_2 - 4x_3 = 0$ y $x_1 + 9x_2 - 4x_3 = -2$.

18. Describa y compare los conjuntos solución *de* $x_1 - 3x_2 + 5x_3 = 0$ y $x_1 - 3x_2 + 5x_3 = 4$.

En los ejercicios 19 y 20, encuentre la ecuación paramétrica de la línea que pasa por **a** y es paralela a **b**.

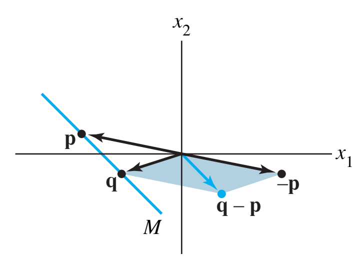
19. $\mathbf{a} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$

20. $\mathbf{a} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -7 \\ 8 \end{bmatrix}$

En los ejercicios 21 y 22, encuentre una ecuación paramétrica de la línea M que pasa por **p** y **q**. [*Indicación:* M es paralela al vector $\mathbf{q} - \mathbf{p}$. Vea la figura que se presenta enseguida.]

21. $\mathbf{p} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$

22. $\mathbf{p} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} 0 \\ -4 \end{bmatrix}$



La línea que pasa por **p** y **q**.

En los ejercicios 23 y 24, señale cada afirmación como verdadera o falsa. Justifique cada respuesta.

23. a. Una ecuación homogénea siempre es consistente.

b. La ecuación $A\mathbf{x} = \mathbf{0}$ proporciona una descripción explícita de su conjunto solución.

c. La ecuación homogénea $A\mathbf{x} = \mathbf{0}$ tiene solución trivial si, y sólo si, cuenta por lo menos con una variable libre.

d. La ecuación $\mathbf{x} = \mathbf{p} + t\mathbf{v}$ describe una línea que pasa por $\mathbf{v}$ paralela a $\mathbf{p}$.

e. El conjunto solución de $A\mathbf{x} = \mathbf{b}$ es el conjunto de todos los vectores de la forma $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, donde $\mathbf{v}_h$ es cualquier solución de la ecuación $A\mathbf{x} = \mathbf{0}$.

24. a. Si $\mathbf{x}$ es una solución no trivial de $A\mathbf{x} = \mathbf{0}$, entonces todas las entradas de $\mathbf{x}$ son diferentes de cero.

b. La ecuación $\mathbf{x} = x_2\mathbf{u} + x_3\mathbf{v}$, con x_2 y x_3 libres (y ni $\mathbf{u}$ ni $\mathbf{v}$ son múltiplos entre sí), describe un plano que pasa por el origen.

c. La ecuación $A\mathbf{x} = \mathbf{b}$ es homogénea si el vector cero es una solución.

d. El efecto de sumar $\mathbf{p}$ a un vector es trasladar el vector en una dirección paralela a $\mathbf{p}$.

e. El conjunto solución de $A\mathbf{x} = \mathbf{b}$ se obtiene al trasladar el conjunto solución de $A\mathbf{x} = \mathbf{0}$.

25. Demuestre el teorema 6:

a. Suponga que $\mathbf{p}$ es una solución de $A\mathbf{x} = \mathbf{b}$, de manera que $A\mathbf{p} = \mathbf{b}$. Sea $\mathbf{v}_h$ cualquier solución de la ecuación homogénea $A\mathbf{x} = \mathbf{0}$, y haga $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$. Muestre que $\mathbf{w}$ es una solución de $A\mathbf{x} = \mathbf{b}$.

b. Sea $\mathbf{w}$ cualquier solución de $A\mathbf{x} = \mathbf{b}$, y defina $\mathbf{v}_h = \mathbf{w} - \mathbf{p}$. Muestre que $\mathbf{v}_h$ es una solución de $A\mathbf{x} = \mathbf{0}$. Esto comprueba que cualquier solución de $A\mathbf{x} = \mathbf{b}$ tiene la forma $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, donde $\mathbf{p}$ es una solución particular de $A\mathbf{x} = \mathbf{b}$ y $\mathbf{v}_h$ es una solución de $A\mathbf{x} = \mathbf{0}$.

26. Suponga que $A\mathbf{x} = \mathbf{b}$ tiene una solución. Explique por qué la solución es única precisamente cuando $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.

27. Suponga que A es la matriz *cero* de 3×3 (con todas las entradas iguales a cero). Describa el conjunto solución de la ecuación $A\mathbf{x} = \mathbf{0}$.

28. Si $\mathbf{b} \neq \mathbf{0}$, ¿el conjunto solución de $A\mathbf{x} = \mathbf{b}$ puede ser un plano que pase por el origen? Explique su respuesta.

En los ejercicios 29 a 32, (a) ¿la ecuación $A\mathbf{x} = \mathbf{0}$ tiene una solución no trivial?, y (b) ¿la ecuación $A\mathbf{x} = \mathbf{b}$ tiene por lo menos una solución para toda $\mathbf{b}$ posible?

29. A es una matriz de 3×3 con tres posiciones pivote.

30. A es una matriz de 3×3 con dos posiciones pivote.

31. A es una matriz de 3×2 con dos posiciones pivote.

32. A es una matriz de 2×4 con dos posiciones pivote.

33. Dada $A = \begin{bmatrix} -2 & -6 \\ 7 & 21 \\ -3 & -9 \end{bmatrix}$, encuentre mediante inspección una solución no trivial de $A\mathbf{x} = \mathbf{0}$. [*Sugerencia:* Piense en la ecuación $A\mathbf{x} = \mathbf{0}$ escrita como una ecuación vectorial.]

34. Dada $A = \begin{bmatrix} 4 & -6 \\ -8 & 12 \\ 6 & -9 \end{bmatrix}$, encuentre mediante inspección una una solución no trivial de $A\mathbf{x} = \mathbf{0}$.

35. Construya una matriz A de 3×3, distinta de cero, tal que el vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ sea una solución de $A\mathbf{x} = \mathbf{0}$.

36. Construya una matriz A de 3×3, distinta de cero, tal que el vector $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ sea una solución de $A\mathbf{x} = \mathbf{0}$.

37. Construya una matriz A de 2×2 tal que el conjunto solución de la ecuación $A\mathbf{x} = \mathbf{0}$ esté en la línea en $\mathbb{R}^2$ que pasa por (4, 1) y el origen. Después, encuentre un vector $\mathbf{b}$ en $\mathbb{R}^2$ tal que el conjunto solución de $A\mathbf{x} = \mathbf{b}$ *no* sea una línea en $\mathbb{R}^2$ paralela al conjunto solución de $A\mathbf{x} = \mathbf{0}$. ¿Por qué esto *no* contradice el teorema 6?

38. Suponga que A es una matriz de 3×3 e $\mathbf{y}$ un vector en $\mathbb{R}^3$ tal que la ecuación $A\mathbf{x} = \mathbf{y}$ *no* tiene una solución. ¿Existe un vector $\mathbf{z}$ en $\mathbb{R}^3$ tal que la ecuación $A\mathbf{x} = \mathbf{z}$ tenga una solución única? Analice el planteamiento.

39. Sea A una matriz de $m \times n$ y $\mathbf{u}$ un vector en $\mathbb{R}^n$ que satisfaga la ecuación $A\mathbf{x} = \mathbf{0}$. Muestre que para cualquier escalar c, el vector $c\mathbf{u}$ también satisface $A\mathbf{x} = \mathbf{0}$. [Esto es, muestre que $A(c\mathbf{u}) = \mathbf{0}$.]

40. Sea A una matriz de $m \times n$, y sean $\mathbf{u}$ y $\mathbf{v}$ vectores en $\mathbb{R}^n$ con la propiedad de que $A\mathbf{u} = \mathbf{0}$ y $A\mathbf{v} = \mathbf{0}$. Explique por qué $A(\mathbf{u} + \mathbf{v})$ debe ser igual al vector cero. Después explique por qué $A(c\mathbf{u} + d\mathbf{v}) = \mathbf{0}$ para cada par de escalares c y d.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 1 & 4 & -5 & 0 \\ 2 & -1 & 8 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & -5 & 0 \\ 0 & -9 & 18 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -2 & -1 \end{bmatrix}$$

$$\begin{aligned} x_1 \quad + 3x_3 &= 4 \\ x_2 - 2x_3 &= -1 \end{aligned}$$

Entonces $x_1 = 4 - 3x_3$, $x_2 = -1 + 2x_3$, con x_3 libre. La solución general en forma vectorial paramétrica es

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 - 3x_3 \\ -1 + 2x_3 \\ x_3 \end{bmatrix} = \underset{\mathbf{p}}{\underset{\uparrow}{\begin{bmatrix} 4 \\ -1 \\ 0 \end{bmatrix}}} + x_3 \underset{\mathbf{v}}{\underset{\uparrow}{\begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}}}$$

La intersección de los dos planos es la línea que pasa por **p** en la dirección de **v**.

2. La matriz aumentada $[10 \;\; -3 \;\; -2 \;\; 7]$ es equivalente por filas a $[1 \;\; -.3 \;\; -.2 \;\; .7]$, y la solución general es $x_1 = .7 + .3x_2 + .2x_3$, con x_2 y x_3 libres. Esto es,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .7 + .3x_2 + .2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .7 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} .3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} .2 \\ 0 \\ 1 \end{bmatrix}$$
$$= \mathbf{p} + x_2\mathbf{u} + x_3\mathbf{v}$$

El conjunto solución de la ecuación no homogénea $A\mathbf{x} = \mathbf{b}$ es el plano trasladado $\mathbf{p} + \text{Gen}\{\mathbf{u}, \mathbf{v}\}$, el cual pasa por **p** y es paralelo al conjunto solución de la ecuación homogénea del ejemplo 2.

1.6 APLICACIONES DE LOS SISTEMAS LINEALES

Podría esperarse que un problema de la vida real que involucre álgebra lineal tuviera sólo una solución, o quizá ninguna. El propósito de esta sección es mostrar cómo pueden surgir, de manera natural, sistemas lineales con muchas soluciones. Las aplicaciones que se presentan aquí tienen que ver con economía, química y flujo de redes.

Un sistema homogéneo en economía

El sistema de 500 ecuaciones con 500 variables mencionado en la introducción de este capítulo se conoce ahora como el modelo "de entrada y salida" (o "de producción") de Leontief.[1] En la sección 2.6 se estudiará con más detalle este modelo, cuando se haya visto más teoría y se cuente con una mejor notación. Por ahora, se examinará un "modelo de intercambio" más sencillo, también desarrollado por Leontief.

Suponga que la economía de una nación se divide en muchos sectores, tales como diversas industrias de fabricación, comunicación, entretenimiento y servicio. Suponga también que se conoce el rendimiento total de cada sector para un año y se sabe exactamente cómo se divide este rendimiento, o "se intercambia", entre los otros sectores de la economía. El valor total en moneda (dólares en este caso) del rendimiento de un sector será el **precio** de dicho rendimiento. Leontief demostró el resultado siguiente.

[1]Vea Wassily W. Leontief, "Input–Output Economics", *Scientific American*, octubre de 1951, pp. 15–21.

Existen *precios de equilibrio* que se pueden asignar a los rendimientos totales de los diversos sectores de manera que los ingresos de cada sector balanceen exactamente sus gastos.

El ejemplo siguiente muestra cómo encontrar los precios de equilibrio.

EJEMPLO 1 Suponga que una economía consiste en los sectores de carbón, electricidad y acero, y que el rendimiento de cada sector se distribuye entre los diferentes sectores como en la tabla 1, donde las entradas de una columna representan fracciones de la producción total de un sector.

La segunda columna de la tabla 1, por ejemplo, muestra que la producción total de electricidad se divide como sigue: un 40% de carbón, un 50% de acero, y el restante 10% de electricidad. (El sector eléctrico trata este 10% como un gasto en que incurre para hacer funcionar su negocio.) Ya que debe tomarse en cuenta la producción total, las fracciones decimales de cada columna deben sumar 1.

Los precios (es decir, valores en moneda) de la producción total de los sectores de carbón, electricidad y acero se denotarán como p_C, p_E y p_S, respectivamente. Si es posible, encuentre los precios de equilibrio que permiten a los ingresos de cada sector igualar sus gastos.

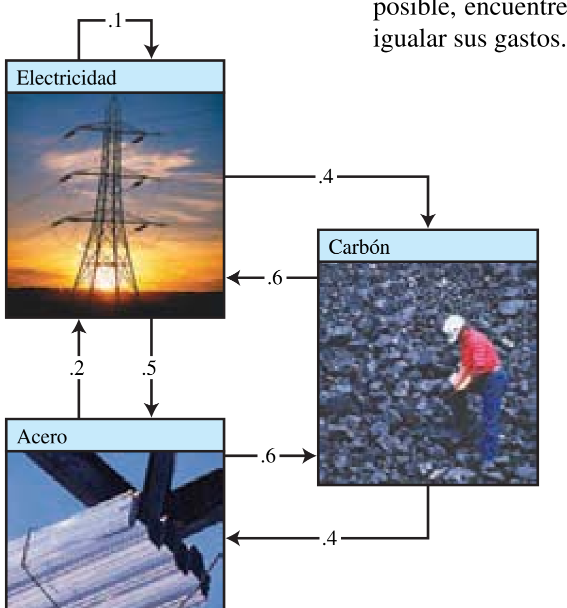


TABLA 1 Una economía sencilla

Distribución del rendimiento de:			
Carbón	Electricidad	Acero	Comprado por:
.0	.4	.6	Carbón
.6	.1	.2	Electricidad
.4	.5	.2	Acero

Solución Un sector observa una columna para ver a dónde va su producción, y examina una fila para ver qué necesita como entradas. Por ejemplo, la primera fila de la tabla 1 indica que el sector carbón recibe (y paga por) el 40% de la producción del sector eléctrico y el 60% de la producción de acero. Puesto que los valores respectivos de producción totales son p_E y p_S, el sector carbón debe gastar $.4p_E$ dólares por su parte de producción de electricidad, y $.6p_S$ por su parte de producción de acero. Entonces los gastos totales del sector carbón son de $.4p_E + .6p_S$. Para hacer que los ingresos del sector carbón, p_C, sean iguales a sus gastos, se desea

$$p_C = .4p_E + .6p_S \tag{1}$$

La segunda fila de la tabla de intercambio muestra que el sector eléctrico gasta $.6p_C$ en carbón, $.1p_E$ en electricidad, y $.2p_S$ en acero. Entonces, el requisito ingreso/gastos

para electricidad es

$$p_E = .6p_C + .1p_E + .2p_S \qquad (2)$$

Por último, la tercera fila de la tabla de intercambio conduce al requisito final:

$$p_S = .4p_C + .5p_E + .2p_S \qquad (3)$$

Para resolver el sistema de ecuaciones (1), (2) y (3), traslade todas las incógnitas al lado izquierdo de las ecuaciones y combínelas como términos. [Por ejemplo, a la izquierda de (2) escriba $p_E - .1p_E$ como $.9p_E$.]

$$\begin{aligned} p_C - .4p_E - .6p_S &= 0 \\ -.6p_C + .9p_E - .2p_S &= 0 \\ -.4p_C - .5p_E + .8p_S &= 0 \end{aligned}$$

Lo que sigue es reducir por filas. Aquí, para simplificar, los decimales se redondean a dos posiciones.

$$\begin{bmatrix} 1 & -.4 & -.6 & 0 \\ -.6 & .9 & -.2 & 0 \\ -.4 & -.5 & .8 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & .66 & -.56 & 0 \\ 0 & -.66 & .56 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & .66 & -.56 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & 1 & -.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -.94 & 0 \\ 0 & 1 & -.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general es $p_C = .94p_S$, $p_E = .85p_S$, y p_S es libre. El vector precio de equilibrio para la economía tiene la forma

$$\mathbf{p} = \begin{bmatrix} p_C \\ p_E \\ p_S \end{bmatrix} = \begin{bmatrix} .94p_S \\ .85p_S \\ p_S \end{bmatrix} = p_S \begin{bmatrix} .94 \\ .85 \\ 1 \end{bmatrix}$$

Cualquier selección (no negativa) para p_S se convierte en una selección de precios de equilibrio. Por ejemplo, si se toma p_S como 100 (o $100 millones), entonces $p_C = 94$ y $p_E = 85$. Los ingresos y gastos de cada sector serán iguales si la producción de carbón se valora en $94 millones, la producción eléctrica en $85 millones, y la producción de acero en $100 millones. ■

Balanceo de ecuaciones químicas

Las ecuaciones químicas describen las cantidades de sustancias consumidas y producidas por las reacciones químicas. Por ejemplo, cuando se quema gas propano (C_3H_8), éste se combina con oxígeno (O_2) para formar dióxido de carbono (CO_2) y agua (H_2O), de acuerdo con una ecuación de la forma

$$(x_1)C_3H_8 + (x_2)O_2 \rightarrow (x_3)CO_2 + (x_4)H_2O \qquad (4)$$

Para "balancear" esta ecuación, un químico debe encontrar números enteros $x_1, \ldots, x_4$ tales que el número total de átomos de carbono (C), hidrógeno (H) y oxígeno (O) situados a la izquierda sea igual al número correspondiente de átomos ubicados a la derecha (porque los átomos no se crean ni se destruyen en la reacción).

Un método sistemático para balancear ecuaciones químicas consiste en establecer una ecuación que describa el número de átomos de cada tipo presente en una reacción. Como la ecuación (4) involucra tres tipos de átomo (carbono, hidrógeno y oxígeno), construya un vector en $\mathbb{R}^3$ para cada reactivo y producto en (4) que enliste el número de “átomos por molécula”, como sigue:

$$C_3H_8: \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix}, \ O_2: \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \ CO_2: \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \ H_2O: \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \begin{matrix} \leftarrow \text{Carbón} \\ \leftarrow \text{Hidrógeno} \\ \leftarrow \text{Oxígeno} \end{matrix}$$

Para balancear la ecuación (4), los coeficientes $x_1, \ldots, x_4$ debe satisfacer

$$x_1\begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = x_3\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + x_4\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

Para resolver, traslade todos los términos a la izquierda (cambiando los signos en los vectores tercero y cuarto):

$$x_1\begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} + x_4\begin{bmatrix} 0 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

La reducción por filas de la matriz aumentada para esta ecuación conduce a la solución general

$$x_1 = \tfrac{1}{4}x_4, \ \ x_2 = \tfrac{5}{4}x_4, \ \ x_3 = \tfrac{3}{4}x_4, \ \ \text{con } x_4 \text{ libre}$$

Como los coeficientes en una ecuación química deben ser enteros, tome $x_4 = 4$, en tal caso, $x_1 = 1$, $x_2 = 5$ y $x_3 = 3$. La ecuación balanceada es

$$C_3H_8 + 5O_2 \rightarrow 3CO_2 + 4H_2O$$

La ecuación también estaría balanceada si, por ejemplo, cada uno de los coeficientes se duplicara. Sin embargo, para la mayoría de los propósitos, los químicos prefieren usar una ecuación balanceada cuyos coeficientes sean los números enteros más pequeños posibles.

Flujo de redes

Los sistemas de ecuaciones lineales surgen de manera natural cuando científicos, ingenieros o economistas estudian el flujo de algunas cantidades a través de una red. Por ejemplo, los planeadores urbanos e ingenieros de tráfico monitorean el patrón de flujo del tráfico en una cuadrícula formada por las calles de una ciudad. Los ingenieros eléctricos calculan el flujo de corriente que transportan los circuitos eléctricos. Y los economistas analizan la distribución de productos entre fabricantes y consumidores que tiene lugar mediante una red de mayoristas y vendedores al menudeo. Para muchas redes, los sistemas de ecuaciones involucran cientos e incluso miles de variables y ecuaciones.

Una *red* consiste en un conjunto de puntos llamados *uniones* o *nodos*, con líneas o arcos denominados *ramas* que conectan a algunos o todos los nodos. La dirección del flujo se indica en cada arco y la cantidad (o tasa) de flujo se muestra o se denota por medio de una variable.

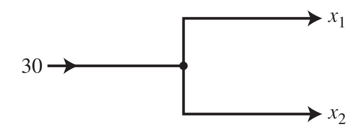


FIGURA 1 Una unión o nodo.

El supuesto básico del flujo de redes es que el flujo que entra a la red es el mismo que sale de la red, y que el flujo entrante en un nodo es igual al flujo saliente del nodo. Por ejemplo, en la figura 1 se muestran 30 unidades que fluyen hacia un nodo a través de un arco, con x_1 y x_2 denotando los flujos que salen del nodo por otros arcos. Como el flujo se "conserva" en cada nodo, debe ser cierto que $x_1 + x_2 = 30$. De manera similar, el flujo en cada nodo se describe por medio de una ecuación lineal. El problema del análisis de redes consiste en determinar el flujo presente en cada arco cuando se conoce cierta información parcial (como las entradas a la red).

EJEMPLO 2 En la red de la figura 2 se muestra el flujo del tráfico (en vehículos por hora) sobre varias calles de un solo sentido en el centro de Baltimore durante una día típico temprano por la tarde. Determine el patrón de flujo general para la red.

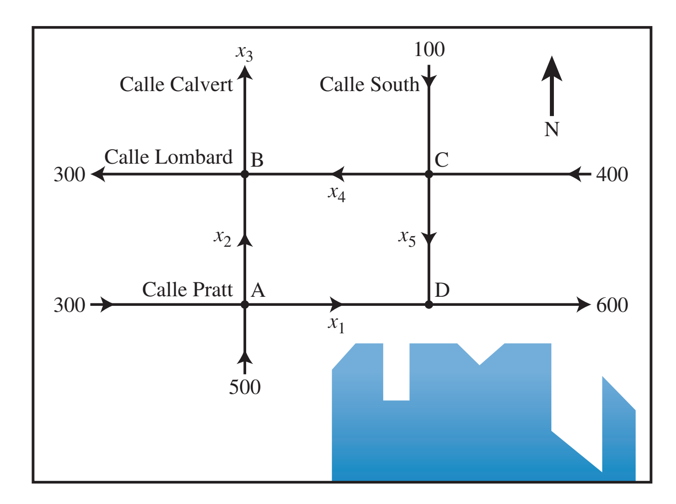


FIGURA 2 Calles de Baltimore.

Solución Anote las ecuaciones que describen el flujo, y después encuentre la solución general del sistema. Etiquete las intersecciones de las calles (nodos) y los flujos desconocidos en los arcos, como se muestra en la figura 2. En cada intersección, establezca el flujo entrante igual al flujo saliente,

Intersección	Flujo entrante		Flujo saliente
A	$300 + 500$	$=$	$x_1 + x_2$
B	$x_2 + x_4$	$=$	$300 + x_3$
C	$100 + 400$	$=$	$x_4 + x_5$
D	$x_1 + x_5$	$=$	600

También, el flujo total entrante a la red $(500 + 300 + 100 + 400)$ es igual al flujo total saliente $(300 + x_3 + 600)$, lo cual se simplifica a $x_3 = 400$. Combine esta ecuación con

un reordenamiento de las primeras cuatro ecuaciones para obtener el siguiente sistema de ecuaciones:

$$\begin{aligned} x_1 + x_2 \qquad\qquad\qquad &= 800 \\ x_2 - x_3 + x_4 \qquad &= 300 \\ x_4 + x_5 &= 500 \\ x_1 \qquad\qquad\qquad + x_5 &= 600 \\ x_3 \qquad\qquad &= 400 \end{aligned}$$

La reducción por filas de la matriz aumentada asociada conduce a

$$\begin{aligned} x_1 \qquad\qquad + x_5 &= 600 \\ x_2 \qquad - x_5 &= 200 \\ x_3 \qquad\quad &= 400 \\ x_4 + x_5 &= 500 \end{aligned}$$

El patrón de flujo general para la red se describe por medio de

$$\begin{cases} x_1 = 600 - x_5 \\ x_2 = 200 + x_5 \\ x_3 = 400 \\ x_4 = 500 - x_5 \\ x_5 \text{ está libre} \end{cases}$$

Un flujo negativo en un arco de red corresponde al flujo que va en dirección opuesta al sentido mostrado en el modelo. Como en este problema las calles van en un solo sentido, ninguna de las variables puede ser negativa. Este hecho conduce a ciertas limitaciones sobre los posibles valores de las variables. Por ejemplo, $x_5 \leq 500$ porque x_4 no puede ser negativa. En el problema de práctica 2 se consideran otras restricciones que actúan sobre las variables. ■

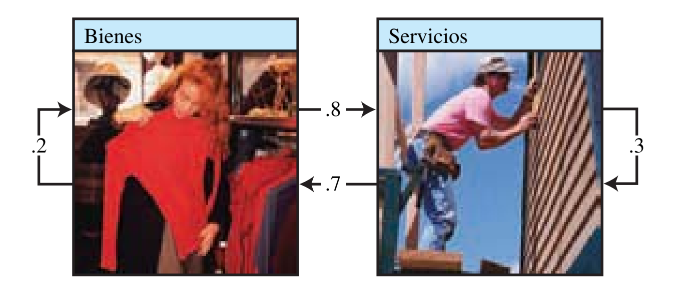
PROBLEMAS DE PRÁCTICA

1. Suponga que una economía tiene tres sectores, agricultura, minería y manufactura. Agricultura vende el 5% de su producción a minería, el 30% a manufactura, y retiene el resto. Minería vende un 20% de lo que produce a agricultura, un 70% a manufactura, y conserva el resto. Manufactura vende el 20% de su producción a agricultura, el 30% a minería, y se queda con el 50%. Determine la tabla de intercambio para esta economía, donde las columnas describen el modo en que la producción de cada sector se intercambia entre los tres sectores.

2. Considere el flujo de la red que se presentó en el ejemplo 2. Determine el rango posible de valores para x_1 y x_2. [*Sugerencia:* En el ejemplo se mostró que $x_5 \leq 500$. ¿Qué implica esto acerca de x_1 y x_2? También utilice el hecho de que $x_5 \geq 0$.]

1.6 Ejercicios

1. Suponga que una economía tiene solamente dos sectores: bienes y servicios. Cada año, bienes vende el 80% de su producción a servicios y se queda con el resto, mientras que servicios vende un 70% de su producción a bienes y retiene el 30%. Para la producción anual de los sectores de bienes y servicios, encuentre precios de equilibrio que permitan que los ingresos de cada sector equivalgan a sus gastos.



2. Encuentre otro conjunto de precios de equilibrio para la economía del ejemplo 1. Suponga que la misma economía usó yenes japoneses en lugar de dólares para medir el valor de la producción en los diferentes sectores. ¿Alteraría esto el problema de alguna manera? Analice este planteamiento.

3. Considere una economía con tres sectores: químicos y metales, combustibles y energía, y maquinaria. Químicos vende el 30% de su producción a combustibles, un 50% a maquinaria, y retiene el resto. Combustibles vende un 80% de su producción a químicos, el 10% a maquinaria, y retiene el 10%. Maquinaria vende el 40% a químicos, el 40% a combustibles y conserva el resto.
   a. Construya la tabla de intercambio para esta economía.
   b. Desarrolle un sistema de ecuaciones que conduzca a precios con los cuales los ingresos de cada sector equivalgan a sus gastos. Luego escriba la matriz aumentada que pueda reducirse por filas para encontrar dichos precios.
   c. **[M]** Encuentre un conjunto de precios de equilibrio cuando el precio para la producción de maquinaria es de 100 unidades.

4. Suponga que una economía tiene cuatro sectores, agricultura (A), energía (E), manufactura (M) y transporte (T). El sector A vende un 10% de su producción a E, el 25% a M, y retiene el resto. El sector E vende un 30% de su producción a A, un 35% a M, un 25% a T, y conserva el resto. El sector M vende el 30% de su producción a A, el 15% a E, un 40% a T, y conserva lo restante. El sector T vende el 20% de su producción a A, el 10% a E, el 30% a M, y se queda con el 40 por ciento.
   a. Construya la tabla de intercambio para esta economía.
   b. **[M]** Encuentre un conjunto de precios de equilibrio para esta economía.

Balancee las ecuaciones químicas de los ejercicios 5 a 10 usando el enfoque con ecuaciones vectoriales que se analizó en esta sección.

5. El sulfato de boro reacciona de manera violenta con el agua para formar ácido bórico y sulfato de hidrógeno gaseoso (el olor de los huevos podridos). La ecuación no balanceada es

   $B_2S_3 + H_2O \to H_3BO_3 + H_2S$

   [Para cada compuesto, construya un vector que enliste el número de átomos de boro, hidrógeno y oxígeno.]

6. Cuando se mezclan soluciones de fosfato de sodio y nitrato de bario, el resultado es fosfato de bario (como un precipitado) y nitrato de sodio. La ecuación no balanceada es

   $Na_3PO_4 + Ba(NO_3)_2 \to Ba_3(PO_4)_2 + NaNO_3$

   [Para cada compuesto, construya un vector que enliste el número de átomos de sodio (Na), fósforo, oxígeno, bario y nitrógeno. Por ejemplo, el nitrato de bario corresponde a (0, 0, 6, 1, 2).]

7. Alka-Seltzer contiene bicarbonato de sodio ($NaHCO_3$) y ácido cítrico ($H_3C_6H_5O_7$). Cuando una tableta se disuelve en agua, la siguiente reacción produce citrato de sodio, agua y dióxido de carbono (gaseoso):

   $NaHCO_3 + H_3C_6H_5O_7 \to Na_3C_6H_5O_7 + H_2O + CO_2$

8. La siguiente reacción entre permanganato de potasio ($KMnO_4$) y sulfato de manganeso en presencia de agua produce dióxido de manganeso, sulfato de potasio y ácido sulfúrico:

   $KMnO_4 + MnSO_4 + H_2O \to MnO_2 + K_2SO_4 + H_2SO_4$

   [Para cada compuesto, construya un vector que enliste el número de átomos de potasio (K), manganeso, oxígeno, azufre e hidrógeno.]

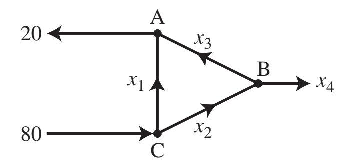
9. **[M]** Si es posible, use aritmética exacta o formato racional para realizar los cálculos necesarios y balancear la siguiente reacción química:

   $PbN_6 + CrMn_2O_8 \to Pb_3O_4 + Cr_2O_3 + MnO_2 + NO$

10. **[M]** La siguiente reacción química puede usarse en algunos procesos industriales, como en la producción de arsénico (AsH_3). Use aritmética exacta o formato racional para realizar los cálculos necesarios y balancear esta ecuación.

   $$MnS + As_2Cr_{10}O_{35} + H_2SO_4 \to HMnO_4 + AsH_3 + CrS_3O_{12} + H_2O$$

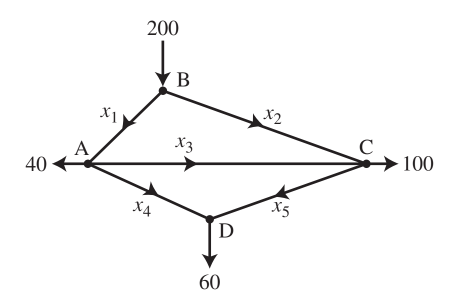
11. Encuentre el patrón de flujo general de la red que se muestra en la figura. Suponiendo que todos los flujos son no negativos, ¿cuál el máximo valor posible para x_3?



12. a. Encuentre el patrón de tráfico general en la red de calles principales que se muestra en la figura. (Las tasas de flujo se dan en automóviles por minuto.)

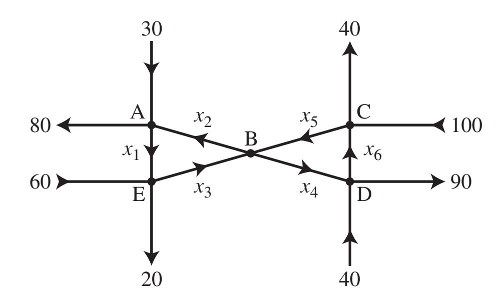
b. Describa el patrón de tráfico general cuando se cierra el camino cuyo flujo es x_4.

c. Cuando $x_4 = 0$, ¿cuál es el valor mínimo de x_1?

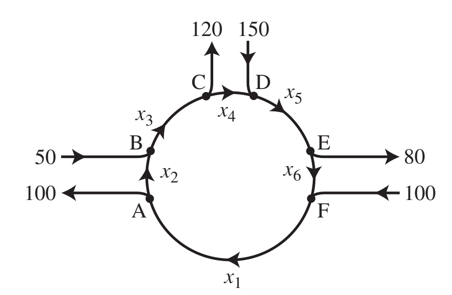


13. a. Encuentre el patrón de flujo general para la red que se muestra en la figura.

b. Si el flujo debe ir en la dirección indicada, ¿cuáles son los flujos mínimos en los arcos denotados por x_2, x_3, x_4 y x_5?



14. A menudo, en Inglaterra las intersecciones se construyen en forma de "glorieta" con un solo sentido, como indica la figura. Suponga que el tráfico debe moverse en la dirección mostrada. Encuentre la solución general del flujo de la red y el mínimo valor posible para x_6.



SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Escriba los porcentajes como decimales. Ya que toda producción debe ser tomada en cuenta, cada una de las columnas ha de sumar 1. Este hecho ayuda a cubrir cualquier entrada faltante.

Distribución de la producción de:			
Agricultura	Minería	Manufactura	Comprada por:
.65	.20	.20	Agricultura
.05	.10	.30	Minería
.30	.70	.50	Manufactura

2. Como $x_1 \leq 500$, las ecuaciones para x_1 y x_2 implican que $x_1 \geq 100$ y $x_2 \leq 700$. El hecho de que $x_5 \geq 0$ implica que $x_1 \leq 600$ y $x_2 \geq 200$. Entonces, $100 \leq x_1 \leq 600$, y $200 \leq x_2 \leq 700$.

1.7 INDEPENDENCIA LINEAL

Las ecuaciones homogéneas de la sección 1.5 pueden estudiarse desde una perspectiva diferente si se escriben como ecuaciones vectoriales. De esta manera, cambia el enfoque de las soluciones desconocidas de $A\mathbf{x} = \mathbf{0}$ a los vectores que aparecen en las ecuaciones vectoriales.

Por ejemplo, considere la ecuación

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{1}$$

Por supuesto, esta ecuación tiene una solución trivial, donde $x_1 = x_2 = x_3 = 0$. Como en la sección 1.5, el aspecto principal a considerar es si la solución trivial es *la única.*

DEFINICIÓN

Un conjunto de vectores indexado $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ es **linealmente independiente** si la ecuación vectorial

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

tiene únicamente la solución trivial. El conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es **linealmente dependiente** si existen pesos $c_1, \ldots, c_p$, no todos iguales a cero, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \tag{2}$$

La ecuación (2) se llama **relación de dependencia lineal** entre $\mathbf{v}_1, \ldots, \mathbf{v}_p$ cuando no todos los pesos son iguales a cero. Un conjunto indexado es linealmente dependiente si no es linealmente independiente. Por brevedad, puede afirmarse que $\mathbf{v}_1, \ldots, \mathbf{v}_p$ son linealmente dependientes cuando se pretenda establecer que $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es un conjunto linealmente dependiente. Se usará una terminología análoga para conjuntos linealmente independientes.

EJEMPLO 1 Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$.

a. Determine si el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente.

b. Si es posible, encuentre una relación de dependencia lineal entre $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Solución

a. Debe determinarse si hay una solución no trivial de la ecuación (1) anterior. Usando operaciones elementales de fila en la matriz aumentada asociada muestre que

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Es claro que x_1 y x_2 son las variables básicas mientras que x_3 es libre. Cada valor diferente de cero de x_3 determina una solución no trivial de (1). Por lo tanto, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ son linealmente dependientes (y no linealmente independientes).

b. Para encontrar una relación de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, realice una reducción por filas completa a la matriz aumentada y escriba el nuevo sistema:

$$\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_1 \qquad - 2x_3 &= 0 \\ x_2 + \ x_3 &= 0 \\ 0 &= 0 \end{aligned}$$

Así, $x_1 = 2x_3$, $x_2 = -x_3$, y x_3 es libre. Seleccione cualquier valor distinto de cero para x_3, por ejemplo $x_3 = 5$. Entonces, $x_1 = 10$ y $x_2 = -5$. Sustituya estos valores en (1) y obtenga

$$10\mathbf{v}_1 - 5\mathbf{v}_2 + 5\mathbf{v}_3 = \mathbf{0}$$

Ésta es una posible relación (existe una infinidad) de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$.

Independencia lineal entre las columnas de una matriz

Suponga que se inicia con una matriz $A = [\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n]$ en lugar de un conjunto de vectores. La ecuación matricial $A\mathbf{x} = \mathbf{0}$ puede escribirse como

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$$

Cada relación de independencia lineal entre las columnas de A corresponde a una solución no trivial de $A\mathbf{x} = \mathbf{0}$. Así, se tiene el siguiente hecho importante.

Las columnas de una matriz A son linealmente independientes si, y sólo si, la ecuación $A\mathbf{x} = \mathbf{0}$ tiene *únicamente* la solución trivial. (3)

EJEMPLO 2 Determine si las columnas de $A = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 2 & -1 \\ 5 & 8 & 0 \end{bmatrix}$ son linealmente independientes.

Solución Para estudiar $A\mathbf{x} = \mathbf{0}$, reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 0 & 1 & 4 & 0 \\ 1 & 2 & -1 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & -2 & 5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 13 & 0 \end{bmatrix}$$

En este punto, es claro que hay tres variables básicas y que no hay variables libres. Por lo tanto, la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial, y las columnas de A son linealmente independientes.

Conjuntos de uno o dos vectores

Un conjunto que contiene sólo un vector —por ejemplo, $\mathbf{v}$— es linealmente independiente si, y sólo si, $\mathbf{v}$ no es el vector cero. Esto se debe a que la ecuación vectorial $x_1\mathbf{v} = \mathbf{0}$ tiene solamente la solución trivial cuando $\mathbf{v} \neq \mathbf{0}$. El vector cero es linealmente dependiente porque $x_1\mathbf{0} = \mathbf{0}$ tiene muchas soluciones no triviales.

El ejemplo siguiente explicará la naturaleza de un conjunto linealmente dependiente de dos vectores.

EJEMPLO 3 Determine si los siguientes conjuntos de vectores son linealmente independientes.

a. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$ b. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$

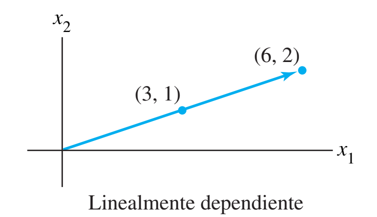
Solución

a. Observe que $\mathbf{v}_2$ es un múltiplo de $\mathbf{v}_1$, a saber, $\mathbf{v}_2 = 2\mathbf{v}_1$. Por lo tanto, $-2\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}$, lo cual muestra que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente dependiente.

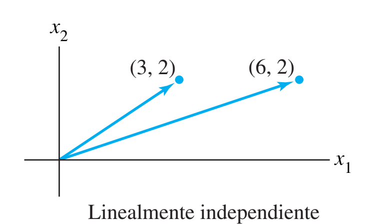
b. Desde luego, los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ *no* son múltiplos el uno del otro. ¿Podrían ser linealmente dependientes? Suponga que c y d satisfacen

$$c\mathbf{v}_1 + d\mathbf{v}_2 = \mathbf{0}$$

Si $c \neq 0$, entonces $\mathbf{v}_1$ puede resolverse en términos de $\mathbf{v}_2$, a saber, $\mathbf{v}_1 = (-d/c)\mathbf{v}_2$. Este resultado es imposible porque $\mathbf{v}_1$ *no* es múltiplo de $\mathbf{v}_2$. Así que c debe ser cero. De manera similar, d debe ser también cero. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto linealmente independiente.



Linealmente dependiente



Linealmente independiente

FIGURA 1

Los argumentos dados en el ejemplo 3 muestran que siempre se puede decidir *por inspección* cuándo un conjunto de dos vectores es linealmente dependiente. No es necesario realizar operaciones de fila. Simplemente verifique si al menos uno de los vectores es un múltiplo escalar del otro. (La prueba es aplicable sólo a conjuntos de *dos* vectores.)

Un conjunto de dos vectores $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente dependiente si, y sólo si, uno de los vectores es múltiplo del otro. El conjunto es linealmente independiente si, y sólo si, ninguno de los vectores es múltiplo del otro.

En términos geométricos, dos vectores son linealmente dependientes si, y sólo si, ambos están sobre la misma línea que pasa por el origen. En la figura 1 se muestran los vectores del ejemplo 3.

Conjuntos de dos o más vectores

La demostración del teorema siguiente es similar a la solución del ejemplo 3. Al final de esta sección se dan los detalles.

TEOREMA 7 **Caracterización de los conjuntos linealmente dependientes**

Un conjunto indexado $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ de dos o más vectores es linealmente dependiente si, y sólo si, al menos uno de los vectores presentes en S es una combinación lineal de los otros. De hecho, si S es linealmente dependiente y $\mathbf{v}_1 \neq \mathbf{0}$, entonces algún $\mathbf{v}_j$ (con $j > 1$) es una combinación lineal de los vectores precedentes, $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$.

Advertencia: El teorema 7 *no* afirma que *cada* vector de un conjunto linealmente dependiente sea una combinación lineal de los vectores precedentes. Un vector de un conjunto linealmente dependiente puede no ser combinación lineal de los otros vectores. Vea el problema de práctica 3.

EJEMPLO 4 Sea $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Describa el conjunto generado por $\mathbf{u}$ y $\mathbf{v}$, y explique por qué un vector $\mathbf{w}$ está en Gen$\{\mathbf{u}, \mathbf{v}\}$ si, y sólo si, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente dependiente.

Solución Los vectores $\mathbf{u}$ y $\mathbf{v}$ son linealmente independientes porque ninguno es múltiplo del otro, así que generan un plano en $\mathbb{R}^3$. (Vea la sección 1.3.) De hecho, Gen$\{\mathbf{u}, \mathbf{v}\}$ es el plano x_1x_2 (con $x_3 = 0$). Si $\mathbf{w}$ es una combinación lineal de $\mathbf{u}$ y de $\mathbf{v}$, entonces $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente dependiente, de acuerdo con el teorema 7. Por otra parte, suponga que $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente dependiente. Por el teorema 7, algún vector en $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es una combinación lineal de los vectores anteriores (puesto que $\mathbf{u} \neq \mathbf{0}$). Ese vector debe ser $\mathbf{w}$, ya que $\mathbf{v}$ no es múltiplo de $\mathbf{u}$. Por lo tanto, $\mathbf{w}$ está en Gen$\{\mathbf{u}, \mathbf{v}\}$. Vea la figura 2.

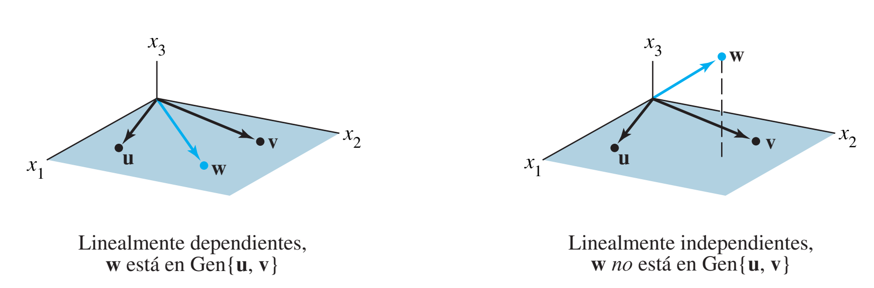


FIGURA 2 Dependencia lineal en $\mathbb{R}^3$.

El ejemplo 4 se generaliza a cualquier conjunto $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ en $\mathbb{R}^3$ con $\mathbf{u}$ y $\mathbf{v}$ linealmente independientes. El conjunto $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ será linealmente dependiente si, y sólo si, $\mathbf{w}$ está en el plano generado por $\mathbf{u}$ y $\mathbf{v}$.

Los siguientes dos teoremas describen casos especiales para los que la dependencia lineal de un conjunto es automática. Además, el teorema 8 resultará clave para efectuar el trabajo en capítulos posteriores.

TEOREMA 8

Si un conjunto contiene más vectores que entradas en cada vector, entonces es linealmente dependiente. Esto es, cualquier conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^3$ es linealmente dependiente si $p > n$.

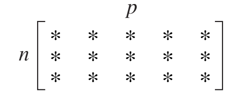


FIGURA 3
Si $p > n$, las columnas son linealmente dependientes.

DEMOSTRACIÓN Sea $A = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_p]$. Entonces A es $n \times p$, y la ecuación $A\mathbf{x} = \mathbf{0}$ corresponde a un sistema de n ecuaciones con p incógnitas. Si $p > n$, hay más variables que ecuaciones, así que debe haber una variable libre. Por lo tanto, $A\mathbf{x} = \mathbf{0}$ tiene una solución no trivial, y las columnas de A son linealmente dependientes. En la figura 3 se muestra una versión matricial de este teorema. ■

Advertencia: El teorema 8 no dice nada acerca del caso en que el número de vectores del conjunto no excede al número de entradas en cada vector.

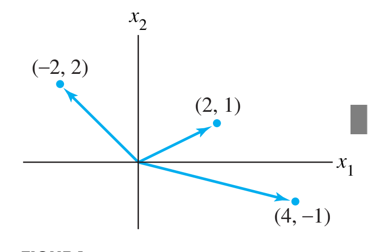


FIGURA 4
Un conjunto linealmente dependiente en $\mathbb{R}^2$.

EJEMPLO 5 Los vectores $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ son linealmente dependientes según el teorema 8, debido a que hay tres vectores en el conjunto y sólo dos entradas en cada vector. Sin embargo, observe que ninguno de los vectores es múltiplo de alguno de los otros vectores. Vea la figura 4. ■

TEOREMA 9

Si un conjunto $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ contiene el vector cero, entonces el conjunto es linealmente dependiente.

DEMOSTRACIÓN Al reordenar los vectores, puede suponerse que $\mathbf{v}_1 = \mathbf{0}$. Entonces la ecuación $1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p = \mathbf{0}$ muestra que S es linealmente dependiente. ■

EJEMPLO 6 Determine por inspección si el conjunto dado es linealmente dependiente.

a. $\begin{bmatrix} 1 \\ 7 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 8 \end{bmatrix}$ b. $\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 8 \end{bmatrix}$ c. $\begin{bmatrix} -2 \\ 4 \\ 6 \\ 10 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \\ -9 \\ 15 \end{bmatrix}$

Solución

a. El conjunto contiene cuatro vectores, cada uno de los cuales tiene sólo tres entradas. Así, el conjunto es linealmente dependiente por el teorema 8.

b. El teorema 8 no aplica aquí porque el número de vectores no excede el número de entradas en cada vector. Como el vector cero pertenece al conjunto, el conjunto es linealmente dependiente por el teorema 9.

c. Compare las entradas correspondientes de los dos vectores. El segundo vector parece ser $-3/2$ veces el primer vector. Esta relación es válida para los primeros tres pares de

entradas, pero no para el cuarto par. Así, ninguno de los vectores es múltiplo del otro y, por lo tanto, son linealmente independientes. ■

SG Dominio de la independencia lineal 1 a 33 (Mastering: Linear Independence 1-33)

En general, se recomienda leer concienzudamente una sección *varias* veces para asimilar un concepto importante como el de independencia lineal. Por ejemplo, la siguiente demostración merece una lectura cuidadosa porque enseña cómo se puede *usar* la relación de independencia lineal.

DEMOSTRACIÓN DEL TEOREMA 7 (Caracterización de conjuntos linealmente dependientes) Si alguna $\mathbf{v}_j$ en S es una combinación lineal de los otros vectores, entonces $\mathbf{v}_j$ puede restarse a cada miembro de la ecuación para producir una relación de dependencia lineal con un peso distinto de cero (-1) en $\mathbf{v}_j$. [Por ejemplo, si $\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, entonces $\mathbf{0} = (-1)\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + 0\mathbf{v}_4 + \cdots + 0\mathbf{v}_p$.] Así que S es linealmente dependiente.

Por otro lado, suponga que S es linealmente dependiente. Si $\mathbf{v}_1$ es cero, entonces es una combinación lineal (trivial) de los otros vectores que hay en S. En caso contrario, $\mathbf{v}_1 \neq \mathbf{0}$, existen pesos $c_1, \ldots, c_p$, no todos cero, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

Sea j el subíndice máximo para el que $c_j \neq 0$. Si $j = 1$, entonces $c_1\mathbf{v}_1 = \mathbf{0}$, lo cual es imposible porque $\mathbf{v}_1 \neq \mathbf{0}$. Así $j > 1$, y

$$c_1\mathbf{v}_1 + \cdots + c_j\mathbf{v}_j + 0\mathbf{v}_{j+1} + \cdots + 0\mathbf{v}_p = \mathbf{0}$$
$$c_j\mathbf{v}_j = -c_1\mathbf{v}_1 - \cdots - c_{j-1}\mathbf{v}_{j-1}$$
$$\mathbf{v}_j = \left(-\frac{c_1}{c_j}\right)\mathbf{v}_1 + \cdots + \left(-\frac{c_{j-1}}{c_j}\right)\mathbf{v}_{j-1}$$ ■

PROBLEMAS DE PRÁCTICA

Sean $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -6 \\ 1 \\ 7 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}$, y $\mathbf{z} = \begin{bmatrix} 3 \\ 7 \\ -5 \end{bmatrix}$.

1. ¿Son los conjuntos $\{\mathbf{u}, \mathbf{v}\}$, $\{\mathbf{u}, \mathbf{w}\}$, $\{\mathbf{u}, \mathbf{z}\}$, $\{\mathbf{v}, \mathbf{w}\}$, $\{\mathbf{v}, \mathbf{z}\}$ y $\{\mathbf{w}, \mathbf{z}\}$ linealmente independientes? ¿Por qué sí o por qué no?
2. ¿La respuesta al problema 1 implica que $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ es linealmente independiente?
3. Para determinar si $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ es linealmente dependiente, es prudente verificar si, por ejemplo, $\mathbf{w}$ es una combinación lineal de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{z}$?
4. ¿$\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ es linealmente dependiente?

1.7 EJERCICIOS

En los ejercicios 1 a 4, determine si los vectores son linealmente independientes. Justifique cada una de sus respuestas.

1. $\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ 2 \\ -6 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -8 \end{bmatrix}$

2. $\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -8 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} -3 \\ 9 \end{bmatrix}$

4. $\begin{bmatrix} -1 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ -8 \end{bmatrix}$

En los ejercicios 5 a 8, determine si las columnas de la matriz dada forman un conjunto linealmente independiente. Justifique cada una de sus respuestas.

5. $\begin{bmatrix} 0 & -8 & 5 \\ 3 & -7 & 4 \\ -1 & 5 & -4 \\ 1 & -3 & 2 \end{bmatrix}$

6. $\begin{bmatrix} -4 & -3 & 0 \\ 0 & -1 & 4 \\ 1 & 0 & 3 \\ 5 & 4 & 6 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 4 & -3 & 0 \\ -2 & -7 & 5 & 1 \\ -4 & -5 & 7 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 1 & -3 & 3 & -2 \\ -3 & 7 & -1 & 2 \\ 0 & 1 & -4 & 3 \end{bmatrix}$

En los ejercicios 9 y 10, (a) ¿para cuáles valores de h está $\mathbf{v}_3$ en Gen$\{\mathbf{v}_1, \mathbf{v}_2\}$?, y (b) ¿para qué valores de h es $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ linealmente *dependiente*? Justifique cada una de sus respuestas.

9. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3 \\ 9 \\ -6 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 5 \\ -7 \\ h \end{bmatrix}$

10. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -5 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -2 \\ 10 \\ 6 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -9 \\ h \end{bmatrix}$

En los ejercicios 11 a 14, encuentre el o los valores de h para los cuales los vectores son linealmente *dependientes*. Justifique cada una de sus respuestas.

11. $\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ h \end{bmatrix}$

12. $\begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix}, \begin{bmatrix} -6 \\ 7 \\ -3 \end{bmatrix}, \begin{bmatrix} 8 \\ h \\ 4 \end{bmatrix}$

13. $\begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ -9 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ h \\ -9 \end{bmatrix}$

14. $\begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} -5 \\ 7 \\ 8 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ h \end{bmatrix}$

Determine por inspección si los vectores en los ejercicios 15 a 20 son linealmente *independientes*. Justifique cada una de sus respuestas.

15. $\begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 8 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 7 \end{bmatrix}$

16. $\begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix}, \begin{bmatrix} 6 \\ -3 \\ 9 \end{bmatrix}$

17. $\begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -6 \\ 5 \\ 4 \end{bmatrix}$

18. $\begin{bmatrix} 4 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 8 \\ 1 \end{bmatrix}$

19. $\begin{bmatrix} -8 \\ 12 \\ -4 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}$

20. $\begin{bmatrix} 1 \\ 4 \\ -7 \end{bmatrix}, \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

En los ejercicios 21 y 22, señale cada enunciado como verdadero o falso. Justifique cada respuesta con base en una lectura cuidadosa del texto.

21. a. Las columnas de una matriz A son linealmente independientes si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene la solución trivial.

b. Si S es un conjunto linealmente dependiente, entonces cada vector es una combinación lineal de los otros vectores en S.

c. Las columnas de cualquier matriz de 4×5 son linealmente dependientes.

d. Si $\mathbf{x}$ e $\mathbf{y}$ son linealmente independientes, y si $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ es linealmente dependiente, entonces $\mathbf{z}$ está en Gen$\{\mathbf{x}, \mathbf{y}\}$.

22. a. Dos vectores son linealmente dependientes si, y sólo si, están en una misma recta que pasa por el origen.

b. Si un conjunto contiene menos vectores que entradas en los vectores, entonces es linealmente independiente.

c. Si $\mathbf{x}$ e $\mathbf{y}$ son linealmente independientes y $\mathbf{z}$ está en Gen$\{\mathbf{x}, \mathbf{y}\}$, entonces $\{\mathbf{x}, \mathbf{y}, \mathbf{z})$ es linealmente dependiente.

d. Si un conjunto en $\mathbb{R}^n$ es linealmente dependiente, entonces el conjunto contiene más vectores que entradas en cada vector.

En los ejercicios 23 a 26, describa las posibles formas escalonadas de la matriz. Utilice la notación del ejemplo 1 dada en la sección 1.2.

23. A es una matriz de 3×3 con columnas linealmente independientes.

24. A es una matriz de 2×2 con columnas linealmente dependientes.

25. A es una matriz de 4×2, $A = [\mathbf{a}_1 \ \ \mathbf{a}_2]$, y $\mathbf{a}_2$ no es múltiplo de $\mathbf{a}_1$.

26. A es una matriz de 4×3, $A = [\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{a}_3]$, tal que $\{\mathbf{a}_1 \ \ \mathbf{a}_2\}$ es linealmente independiente y $\mathbf{a}_3$ no está en Gen$\{\mathbf{a}_1 \ \ \mathbf{a}_2\}$.

27. ¿Cuántas columnas pivote debe tener una matriz de 7×5 si sus columnas son linealmente independientes?¿Por qué?

28. ¿Cuántas columnas pivote debe tener una matriz de 5×7 si sus columnas generan a $\mathbb{R}^5$?¿Por qué?

29. Construya dos matrices A y B de 3×2 tales que $A\mathbf{x} = \mathbf{0}$ tenga únicamente la solución trivial, y $B\mathbf{x} = \mathbf{0}$ tenga una solución no trivial.

30. a. Llene el espacio en blanco de la siguiente afirmación: "Si A es una matriz de $m \times n$, entonces las columnas de A son linealmente independientes si, y sólo si, A tiene _______ columnas pivote".

b. Explique por qué la afirmación en (a) es verdadera.

Los ejercicios 31 y 32 deben resolverse *sin realizar operaciones de fila.* [*Sugerencia:* Escriba $A\mathbf{x} = \mathbf{0}$ como una ecuación vectorial.]

31. Dada $A = \begin{bmatrix} 2 & 3 & 5 \\ -5 & 1 & -4 \\ -3 & -1 & -4 \\ 1 & 0 & 1 \end{bmatrix}$, observe que la tercera columna es la suma de las dos primeras columnas. Encuentre una solución no trivial de $A\mathbf{x} = \mathbf{0}$.

32. Dada $A = \begin{bmatrix} 4 & 1 & 6 \\ -7 & 5 & 3 \\ 9 & -3 & 3 \end{bmatrix}$, observe que la primera columna más dos veces la segunda es igual a la tercera. Encuentre una solución no trivial de $A\mathbf{x} = \mathbf{0}$.

Cada enunciado de los ejercicios 33 a 38 es verdadero (en todos los casos) o bien falso (para al menos un ejemplo). Si la afirmación es falsa, proporcione un ejemplo específico donde muestre que el enunciado no siempre es cierto. Tal ejemplo se llama *contraejemplo* del enunciado. Si la afirmación es cierta, formule una justificación. (Un ejemplo específico no puede explicar por qué una afirmación siempre es cierta. Se tendrá que trabajar más aquí que en los ejercicios 21 y 22.)

33. Si $\mathbf{v}_1, \ldots, \mathbf{v}_4$ están en $\mathbb{R}^4$ y $\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es linealmente dependiente.

34. Si $\mathbf{v}_1, \ldots, \mathbf{v}_4$ están en $\mathbb{R}^4$ y $\mathbf{v}_3 = \mathbf{0}$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es linealmente dependiente.

35. Si $\mathbf{v}_1$ y $\mathbf{v}_2$ están en $\mathbb{R}^4$ y $\mathbf{v}_2$ no es un múltiplo escalar de $\mathbf{v}_1$, entonces $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente independiente.

36. Si $\mathbf{v}_1, \ldots, \mathbf{v}_4$ están en $\mathbb{R}^4$ y $\mathbf{v}_3$ *no* es una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es linealmente independiente.

37. Si $\mathbf{v}_1, \ldots, \mathbf{v}_4$ están en $\mathbb{R}^4$ y $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ también es linealmente dependiente.

38. Si $\mathbf{v}_1, \ldots, \mathbf{v}_4$ son vectores linealmente independientes en $\mathbb{R}^4$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ también es linealmente independiente. [*Sugerencia:* Piense en $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + 0 \cdot \mathbf{v}_4 = \mathbf{0}$.]

39. Suponga que A es una matriz de $m \times n$ con la propiedad de que para cada $\mathbf{b}$ en $\mathbb{R}^m$ la ecuación $A\mathbf{x} = \mathbf{b}$ tiene cuando mucho una solución. Utilice la definición de independencia lineal para explicar por qué las columnas de A deben de ser linealmente independientes.

40. Suponga que una matriz A de $m \times n$ tiene n columnas pivote. Explique por qué para cada $\mathbf{b}$ en $\mathbb{R}^m$ la ecuación $A\mathbf{x} = \mathbf{b}$ tiene cuando mucho una solución. [*Indicación:* Explique por qué $A\mathbf{x} = \mathbf{b}$ no puede tener infinidad de soluciones.]

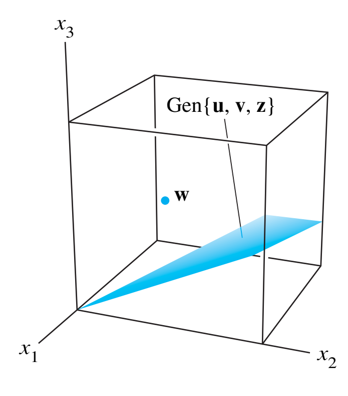
[M] En los ejercicios 41 y 42, use tantas columnas de A como sea posible para construir una matriz B con la propiedad de que la ecuación $B\mathbf{x} = \mathbf{0}$ tenga solamente la solución trivial. Resuelva $B\mathbf{x} = \mathbf{0}$ para verificar su trabajo.

41. $A = \begin{bmatrix} 8 & -3 & 0 & -7 & 2 \\ -9 & 4 & 5 & 11 & -7 \\ 6 & -2 & 2 & -4 & 4 \\ 5 & -1 & 7 & 0 & 10 \end{bmatrix}$

42. $A = \begin{bmatrix} 12 & 10 & -6 & -3 & 7 & 10 \\ -7 & -6 & 4 & 7 & -9 & 5 \\ 9 & 9 & -9 & -5 & 5 & -1 \\ -4 & -3 & 1 & 6 & -8 & 9 \\ 8 & 7 & -5 & -9 & 11 & -8 \end{bmatrix}$

43. **[M]** Con A y B como las del ejercicio 41, elija una columna $\mathbf{v}$ de A que no se haya usado en la construcción de B, y determine si $\mathbf{v}$ está en el conjunto generado por las columnas de B. (Describa sus cálculos.)

44. **[M]** Repita el ejercicio 43 con las matrices A y B del ejercicio 42. Después proporcione una explicación de lo que descubra, suponiendo que B se construyó de la manera especificada.



SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sí. En cada caso, ningún vector es múltiplo del otro. Por lo tanto, cada conjunto es linealmente independiente.
2. No. La observación que aparece en el problema de práctica número 1 no dice nada, por sí sola, acerca de la independencia lineal de $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$.
3. No. Cuando se prueba la independencia lineal, normalmente no es recomendable verificar si un vector elegido es combinación lineal de los otros vectores. Podría su-

ceder que el vector seleccionado no fuera una combinación lineal de los demás y, aun así, el conjunto completo de vectores resultara ser linealmente dependiente. En este problema de práctica, **w** no es combinación lineal de **u**, **v** y **z**.

4. Sí, por el teorema 8. Existen más vectores (cuatro) que entradas (tres) en ellos.

1.8 INTRODUCCIÓN A LAS TRANSFORMACIONES LINEALES

La diferencia entre una ecuación matricial $A\mathbf{x} = \mathbf{b}$ y la ecuación vectorial asociada $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ es sólo un asunto de notación. Sin embargo, una ecuación matricial $A\mathbf{x} = \mathbf{b}$ puede aparecer en álgebra lineal (y en aplicaciones como la graficación por computadora y el procesamiento de señales) en una manera que no esté directamente relacionada con combinaciones lineales de vectores. Esto sucede cuando se piensa en la matriz A como un objeto que "actúa" sobre un vector **x** multiplicándolo para producir un nuevo vector llamado $A\mathbf{x}$.

Por ejemplo, las ecuaciones

$$\underset{\uparrow \atop A}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}} \underset{\uparrow \atop \mathbf{x}}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}} = \underset{\uparrow \atop \mathbf{b}}{\begin{bmatrix} 5 \\ 8 \end{bmatrix}} \quad \text{y} \quad \underset{\uparrow \atop A}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}} \underset{\uparrow \atop \mathbf{u}}{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}} = \underset{\uparrow \atop \mathbf{0}}{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

establecen que multiplicar por A transforma a **x** en **b** y a **u** en el vector cero. Vea la figura 1.

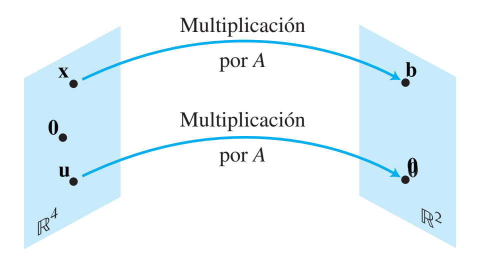


FIGURA 1 Transformación de vectores mediante multiplicación de matrices.

Desde este nuevo punto de vista, la resolución de la ecuación $A\mathbf{x} = \mathbf{b}$ equivale a encontrar todos los vectores **x** en $\mathbb{R}^4$ que se transformen en el vector **b** en $\mathbb{R}^2$ bajo la "acción" que representa multiplicar por A.

La correspondencia de **x** a $A\mathbf{x}$ se denomina *función* de un conjunto de vectores a otro. Este concepto generaliza el conocimiento usual de función como una regla que transforma un número real en otro.

Una **transformación** (o **función** o **mapeo**) T de $\mathbb{R}^n$ a $\mathbb{R}^m$ es una regla que asigna a cada vector **x** en $\mathbb{R}^n$ un vector $T(\mathbf{x})$ en $\mathbb{R}^m$. El conjunto $\mathbb{R}^n$ se llama **dominio** de T, y $\mathbb{R}^m$ se

llama **codominio** de T. La notación $T : \mathbb{R}^n \to \mathbb{R}^m$ indica que el dominio de T es $\mathbb{R}^n$ y que el codominio es $\mathbb{R}^m$. Para $\mathbf{x}$ en $\mathbb{R}^n$, el vector $T(\mathbf{x})$ en $\mathbb{R}^m$ se denomina **imagen** de $\mathbf{x}$ (bajo la acción de T). El conjunto de todas las imágenes $T(\mathbf{x})$ es llamado **rango** de T.

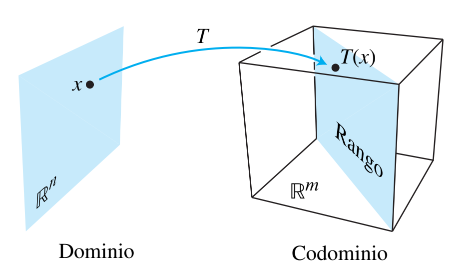
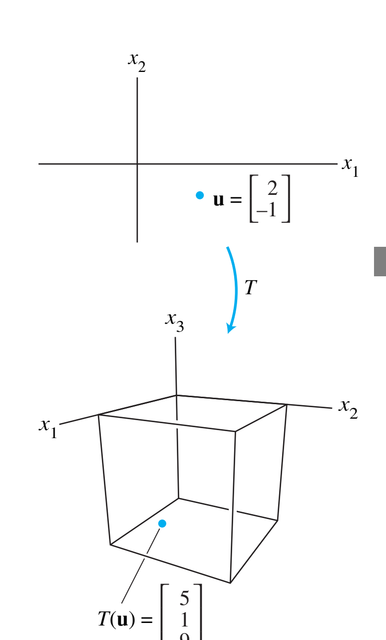


FIGURA 2 Dominio, codominio y rango de $T : \mathbb{R}^n \to \mathbb{R}^m$.

La nueva terminología presentada en esta sección es importante porque la visión dinámica de la multiplicación por matrices es clave para entender muchas ideas del álgebra lineal y estructurar modelos matemáticos de sistemas físicos que evolucionan a través del tiempo. Estos *sistemas dinámicos* se analizarán en las secciones 1.10, 4.8 y 4.9, y a lo largo del capítulo 5.

Transformaciones matriciales

El resto de esta sección se centra en mapeos asociados con la multiplicación de matrices. Para cada $\mathbf{x}$ en $\mathbb{R}^n$, $T(\mathbf{x})$ se calcula como $A\mathbf{x}$, donde A es una matriz de $m \times n$. En aras de la sencillez, algunas veces tales *transformaciones matriciales* se denotarán mediante $\mathbf{x} \mapsto A\mathbf{x}$. Observe que el dominio de T es $\mathbb{R}^n$ cuando A tiene n columnas y el codominio de T es $\mathbb{R}^m$ cuando cada columna de A tiene m entradas. El rango de T es el conjunto de todas las combinaciones lineales de las columnas de A, porque cada imagen $T(\mathbf{x})$ es de la forma $A\mathbf{x}$.



EJEMPLO 1 Sea $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$, y defina una transformación $T : \mathbb{R}^2 \to \mathbb{R}^3$ por medio de $T(\mathbf{x}) = A\mathbf{x}$, tal que

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

a. Encuentre $T(\mathbf{u})$, la imagen de $\mathbf{u}$ bajo la transformación T.
b. Encuentre una $\mathbf{x}$ en $\mathbb{R}^2$ cuya imagen bajo T sea $\mathbf{b}$.
c. ¿Existe más de una $\mathbf{x}$ cuya imagen bajo T sea $\mathbf{b}$?
d. Determine si $\mathbf{c}$ está en el rango de la transformación T.

Solución

a. Calcule

$$T(\mathbf{u}) = A\mathbf{u} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix} \tag{1}$$

b. Resuelva $T(\mathbf{x}) = \mathbf{b}$ para $\mathbf{x}$. Esto es, resuelva $A\mathbf{x} = \mathbf{b}$, o bien

$$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$$

Usando el método de la sección 1.4, reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \tag{2}$$

Por lo tanto, $x_1 = 1.5$, $x_2 = -.5$, y $\mathbf{x} = \begin{bmatrix} 1.5 \\ -.5 \end{bmatrix}$. La imagen de esta $\mathbf{x}$ bajo T es el vector $\mathbf{b}$ dado.

c. Cualquier $\mathbf{x}$ cuya imagen bajo T sea $\mathbf{b}$ debe satisfacer (1). A partir de (2), queda claro que la ecuación (1) tiene una solución única. Así que existe exactamente una $\mathbf{x}$ cuya imagen es $\mathbf{b}$.

d. El vector $\mathbf{c}$ está en el rango de T si $\mathbf{c}$ es la imagen de alguna $\mathbf{x}$ en $\mathbb{R}^2$, esto es, si $\mathbf{c} = T(\mathbf{x})$ para alguna $\mathbf{x}$. Ésta es sólo otra manera de preguntarse si el sistema $A\mathbf{x} = \mathbf{c}$ es consistente. Para encontrar la respuesta, reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 14 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}$$

La tercera ecuación, $0 = -35$, muestra que el sistema es inconsistente. Por lo tanto, $\mathbf{c}$ *no* está en el rango de T.

La pregunta del ejemplo l(c) es un problema de *unicidad* para un sistema de ecuaciones lineales, traducida al idioma de las transformaciones matriciales: ¿Es $\mathbf{b}$ la imagen de una $\mathbf{x}$ *única* en $\mathbb{R}^n$? De manera similar, el ejemplo l(d) es un problema de *existencia*: ¿*Existe* una $\mathbf{x}$ cuya imagen sea $\mathbf{c}$?

Las siguientes dos transformaciones matriciales pueden visualizarse en forma geométrica. Refuerzan la visión dinámica de que una matriz es algo que transforma vectores en otros vectores. La sección 2.7 contiene otros ejemplos interesantes relacionados con la graficación por computadora.

EJEMPLO 2 Si $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ *proyecta* puntos en $\mathbb{R}^3$ sobre el plano x_1x_2 porque

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

Vea la figura 3.

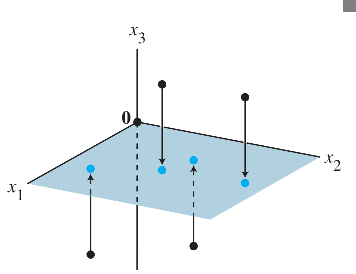
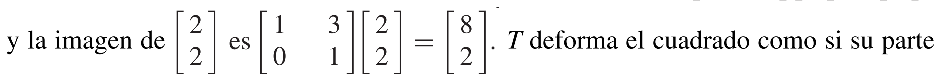
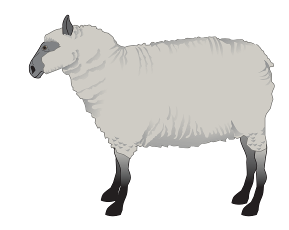


FIGURA 3
Una transformación de proyección.

EJEMPLO 3 Sea $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$. La transformación $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(\mathbf{x}) = A\mathbf{x}$ se llama **transformación de trasquilado**. Es posible demostrar que si T actúa en cada punto del cuadrado de 2×2 que se muestra en la figura 4, entonces el conjunto de imágenes forma el paralelogramo sombreado. La idea central es demostrar que T mapea segmentos de línea sobre segmentos de línea (como se muestra en el ejercicio 27) y comprobar luego que las esquinas del cuadrado se mapean sobre los vértices del paralelogramo. Por ejemplo, la imagen del punto $\mathbf{u} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ es $T(\mathbf{u}) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$,



y la imagen de $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ es $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$. T deforma el cuadrado como si su parte superior fuera empujada hacia la derecha mientras que la base se mantiene fija. Las transformaciones de trasquilado aparecen en física, geología y cristalografía.



Borrego

Borrego trasquilado

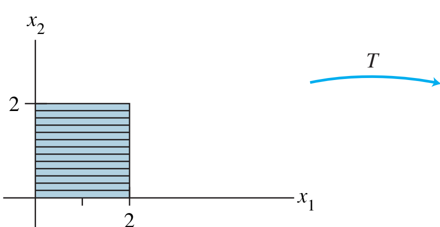


FIGURA 4 Una transformación de trasquilado.

Transformaciones lineales

El teorema 5 de la sección 1.4 muestra que si A es de $m \times n$, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ tiene las propiedades

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} \quad \text{y} \quad A(c\mathbf{u}) = cA\mathbf{u}$$

para cada $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^n$ y todos los escalares c. Estas propiedades, escritas en notación de funciones, identifican la clase más importante de transformaciones del álgebra lineal.

DEFINICIÓN Una transformación (o mapeo) T es **lineal** si:

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para toda $\mathbf{u}$, $\mathbf{v}$ en el dominio de T;

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ para toda $\mathbf{u}$ y todos los escalares c.

Cualquier transformación matricial es una transformación lineal. En los capítulos 4 y 5 se analizarán ejemplos importantes de transformaciones lineales que no son transformaciones matriciales.

Las transformaciones lineales *conservan las operaciones de suma de vectores y multiplicación por un escalar.* La propiedad (1) sostiene que el resultado $T(\mathbf{u} + \mathbf{v})$ sumando primero $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, y aplicando luego T, es el mismo que si primero se aplica T a $\mathbf{u}$ y a $\mathbf{v}$ y luego se suman $T(\mathbf{u})$ y $T(\mathbf{v})$ en $\mathbb{R}^m$. Estas dos propiedades conducen fácilmente a los útiles fundamentos siguientes.

Si T es una transformación lineal, entonces

$$T(\mathbf{0}) = \mathbf{0} \tag{3}$$

y

$$T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v}) \tag{4}$$

para todos los vectores $\mathbf{u}$, $\mathbf{v}$ en el dominio de T y todos los escalares c, d.

La propiedad (3) se deriva de (ii), porque $T(\mathbf{0}) = T(0\mathbf{u}) = 0T(\mathbf{u}) = \mathbf{0}$. La propiedad (4) requiere tanto de (i) como de (ii):

$$T(c\mathbf{u} + d\mathbf{v}) = T(c\mathbf{u}) + T(d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$$

Observe que *si una transformación satisface* (4) *para todas* $\mathbf{u}$, $\mathbf{v}$, c *y* d, *entonces tiene que ser lineal.* (Se establece $c = d = 1$ para la conservación de la suma, y $d = 0$ para conservar la multiplicación por escalares.) Al aplicar (4) en forma repetida se obtiene una generalización útil:

$$T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p) \tag{5}$$

En física e ingeniería, (5) se denomina *principio de superposición.* Piense en $\mathbf{v}_1, \ldots,$ $\mathbf{v}_p$ como señales que entran en un sistema o proceso, y en $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)$ como las respuestas de ese sistema o proceso a dichas señales. El sistema satisface el principio de superposición si al expresar una entrada como una combinación lineal de tales señales, la respuesta del sistema es *la misma* combinación lineal de respuestas a las señales individuales. Esta idea se abordará de nuevo en el capítulo 4.

EJEMPLO 4 Dado un escalar r, se define $T : \mathbb{R}^2 \to \mathbb{R}^2$ como $T(\mathbf{x}) = r\mathbf{x}$. Se dice que T es una **contracción** cuando $0 \le r \le 1$, y es una **dilatación** cuando $r > 1$. Sea $r = 3$, demuestre que T es una transformación lineal.

Solución Sean **u**, **v** en $\mathbb{R}^2$ y c, d escalares. Entonces,

$$\begin{aligned} T(c\mathbf{u}+d\mathbf{v}) &= 3(c\mathbf{u}+d\mathbf{v}) && \text{Definición de } T \\ &= 3c\mathbf{u}+3d\mathbf{v} && \Big\} \text{ Aritmética vectorial} \\ &= c(3\mathbf{u})+d(3\mathbf{v}) && \\ &= cT(\mathbf{u})+dT(\mathbf{v}) \end{aligned}$$

Por lo tanto, T es una transformación lineal porque satisface (4). Vea la figura 5.

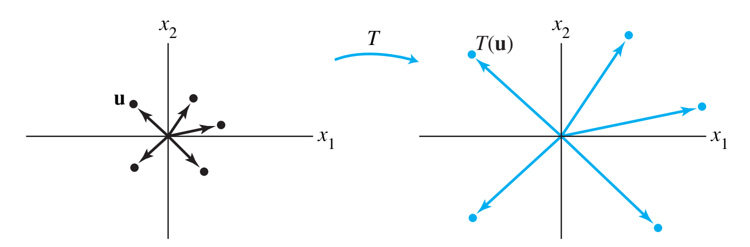


FIGURA 5 Una transformación dilatación.

EJEMPLO 5 Una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ se define como

$$T(\mathbf{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

Encuentre las imágenes bajo T de $\mathbf{u} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, y $\mathbf{u}+\mathbf{v} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$.

Solución

$$T(\mathbf{u}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \qquad T(\mathbf{v}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix},$$

$$T(\mathbf{u}+\mathbf{v}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$

Observe que $T(\mathbf{u}+\mathbf{v})$ es, desde luego, igual a $T(\mathbf{u})+T(\mathbf{v})$. En la figura 6 parece que T hace girar **u**, **v** y $\mathbf{u}+\mathbf{v}$ un ángulo de 90° en sentido contrario al de las manecillas del reloj. De hecho, T transforma el paralelogramo completo determinado por **u** y **v** en otro determinado por $T(\mathbf{u})$ y $T(\mathbf{v})$. (Vea el ejercicio 28.)

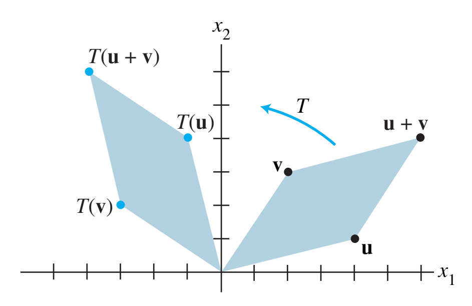


FIGURA 6 Una transformación rotación.

El ejemplo final no es geométrico, sino que muestra cómo un mapeo lineal puede transformar un tipo de datos en otro.

EJEMPLO 6 Una compañía fabrica dos productos, B y C. Usando los datos del ejemplo 7 dados en la sección 1.3, se construye una matriz de “costo unitario”, $U = [\mathbf{b} \quad \mathbf{c}]$, cuyas columnas describen los “costos de producción por dólar” para los distintos productos:

$$\begin{array}{c} \text{Producto} \\ \begin{array}{cc} \text{B} & \text{C} \end{array} \end{array}$$

$$U = \begin{bmatrix} .45 & .40 \\ .25 & .35 \\ .15 & .15 \end{bmatrix} \begin{array}{l} \text{Materiales} \\ \text{Mano de obra} \\ \text{Gastos generales} \end{array}$$

Sea $\mathbf{x} = (x_1, x_2)$ un vector de “producción”, correspondiente a x_1 dólares del producto B y x_2 dólares del producto C, y defina $T : \mathbb{R}^2 \to \mathbb{R}^3$ como

$$T(\mathbf{x}) = U\mathbf{x} = x_1 \begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} + x_2 \begin{bmatrix} .40 \\ .35 \\ .15 \end{bmatrix} = \begin{bmatrix} \text{Costo total de materiales} \\ \text{Costo total de mano de obra} \\ \text{Costo total de gastos generales} \end{bmatrix}$$

El mapeo T transforma una lista de cantidades de producción (medida en dólares, o en otra moneda) en una lista de costos totales. La linealidad de este mapeo se refleja de dos maneras:

1. Si, por ejemplo, la producción se incrementa por un factor de 4, de $\mathbf{x}$ a $4\mathbf{x}$, entonces los costos se incrementarán por el mismo factor, de $T(\mathbf{x})$ a $4T(\mathbf{x})$.
2. Si $\mathbf{x}$ e $\mathbf{y}$ son vectores de producción, entonces el costo total asociado a la producción combinada $\mathbf{x} + \mathbf{y}$ es precisamente la suma de los vectores de costo $T(\mathbf{x})$ y $T(\mathbf{y})$.

PROBLEMAS DE PRÁCTICA

1. Suponga que $T : \mathbb{R}^5 \to \mathbb{R}^2$ y $T(\mathbf{x}) = A\mathbf{x}$ para alguna matriz A y para cada $\mathbf{x}$ en $\mathbb{R}^5$. ¿Cuántas filas y columnas tendrá A?
2. Sea $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Proporcione una descripción geométrica de la transformación $\mathbf{x} \mapsto A\mathbf{x}$.
3. El segmento de recta desde $\mathbf{0}$ hasta un vector $\mathbf{u}$ es el conjunto de puntos de la forma $t\mathbf{u}$, donde $0 \le t \le 1$. Muestre que una transformación lineal T mapea este segmento al segmento que está entre $\mathbf{0}$ y $T(\mathbf{u})$.

1.8 EJERCICIOS

1. Sea $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, y defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ mediante $T(\mathbf{x}) = A\mathbf{x}$. Encuentre las imágenes bajo T de $\mathbf{u} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$.

2. Sea $A = \begin{bmatrix} .5 & 0 & 0 \\ 0 & .5 & 0 \\ 0 & 0 & .5 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix}$, y $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Defina $T : \mathbb{R}^3 \to \mathbb{R}^3$ mediante $T(\mathbf{x}) = A\mathbf{x}$. Encuentre $T(\mathbf{u})$ y $T(\mathbf{v})$.

En los ejercicios 3 a 6, con T definida como $T(\mathbf{x}) = A\mathbf{x}$, encuentre un vector $\mathbf{x}$ cuya imagen bajo T sea $\mathbf{b}$, y determine si esta $\mathbf{x}$ es única.

3. $A = \begin{bmatrix} 1 & 0 & -2 \\ -2 & 1 & 6 \\ 3 & -2 & -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -1 \\ 7 \\ -3 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & -4 \\ 3 & -5 & -9 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 6 \\ -7 \\ -9 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & -5 & -7 \\ -3 & 7 & 5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & -2 & 1 \\ 3 & -4 & 5 \\ 0 & 1 & 1 \\ -3 & 5 & -4 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 9 \\ 3 \\ -6 \end{bmatrix}$

7. Sea A una matriz de 6×5. ¿Cómo deben ser a y b para definir $T : \mathbb{R}^a \to \mathbb{R}^b$ mediante $T(\mathbf{x}) = A\mathbf{x}$?

8. ¿Cuántas filas y columnas debe tener una matriz A para que defina un mapeo de $\mathbb{R}^4$ en $\mathbb{R}^5$ mediante la regla $T(\mathbf{x}) = A\mathbf{x}$?

Para los ejercicios 9 y 10, encuentre todas las $\mathbf{x}$ en $\mathbb{R}^4$ que se mapeen en el vector cero mediante la transformación $\mathbf{x} \mapsto A\mathbf{x}$ para la matriz A dada.

9. $A = \begin{bmatrix} 1 & -4 & 7 & -5 \\ 0 & 1 & -4 & 3 \\ 2 & -6 & 6 & -4 \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & 3 & 9 & 2 \\ 1 & 0 & 3 & -4 \\ 0 & 1 & 2 & 3 \\ -2 & 3 & 0 & 5 \end{bmatrix}$

11. Sea $\mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, y A la matriz del ejercicio 9. ¿Está $\mathbf{b}$ en el rango de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$? ¿Por qué sí o por qué no?

12. Sea $\mathbf{b} = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 4 \end{bmatrix}$, y A la matriz del ejercicio 10. ¿Está $\mathbf{b}$ en el rango de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$? ¿Por qué sí o por qué no?

En los ejercicios 13 a 16, use un sistema de coordenadas rectangulares para graficar $\mathbf{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, y sus imágenes bajo la transformación T dada. (Trace un bosquejo razonablemente grande para cada uno de los ejercicios.) Proporcione una descripción geométrica de lo que T hace a un vector $\mathbf{x}$ en $\mathbb{R}^2$.

13. $T(\mathbf{x}) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

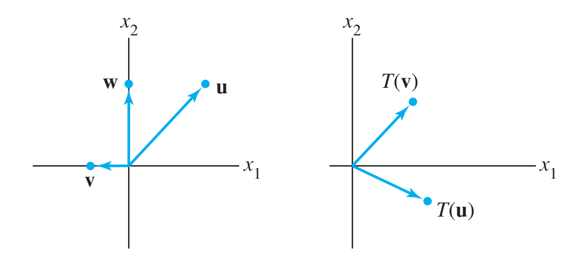
14. $T(\mathbf{x}) = \begin{bmatrix} .5 & 0 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

15. $T(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

16. $T(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

17. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal que mapea $\mathbf{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ en $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ en $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ Use el hecho de que T es lineal para encontrar las imágenes bajo T de $3\mathbf{u}$, $2\mathbf{v}$ y $3\mathbf{u} + 2\mathbf{v}$.

18. La figura muestra los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ junto con las imágenes $T(\mathbf{u})$ y $T(\mathbf{v})$ bajo la acción de una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$. Copie cuidadosamente esta figura, y luego dibuje la imagen $T(\mathbf{w})$ con tanta precisión como sea posible. [*Sugerencia:* Primero, escriba $\mathbf{w}$ como una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$.]



19. Sea $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{y}_1 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, y $\mathbf{y}_2 = \begin{bmatrix} -1 \\ 6 \end{bmatrix}$, y sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal que mapea $\mathbf{e}_1$ en $\mathbf{y}_1$ y $\mathbf{e}_2$ en $\mathbf{y}_2$. Encuentre las imágenes de $\begin{bmatrix} 5 \\ -3 \end{bmatrix}$ y $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

20. Sea $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$, y $\mathbf{v}_2 = \begin{bmatrix} 7 \\ -3 \end{bmatrix}$, y sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal que mapea $\mathbf{x}$ en $x_1\mathbf{v}_1 + x_2\mathbf{v}_2$. Encuentre una matriz tal que $T(\mathbf{x})$ sea $A\mathbf{x}$ para cada $\mathbf{x}$.

En los ejercicios 21 y 22, señale cada enunciado como verdadero o falso. Justifique cada una de sus respuestas.

21. a. Una transformación lineal es un tipo especial de función.

b. Si A es una matriz de 3×5 y T una transformación definida por $T(\mathbf{x}) = A\mathbf{x}$, entonces el dominio de T es $\mathbb{R}^3$.

c. Si A es una matriz de $m \times n$, entonces el rango de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es $\mathbb{R}^2$.

d. Toda transformación lineal es una transformación matricial.

e. Una transformación lineal T es lineal si, y sólo si, $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$ para toda $\mathbf{v}_1$ y $\mathbf{v}_2$ en el dominio de T y para todos los escalares c_1 y c_2.

22. a. Toda transformación matricial es una transformación lineal.

b. El codominio de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es el conjunto de todas las combinaciones lineales de las columnas de A.

c. Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal y $\mathbf{c}$ está en $\mathbb{R}^m$, entonces una pregunta de unicidad es: "¿Está $\mathbf{c}$ en el rango de T?"

d. Una transformación lineal conserva las operaciones de suma de vectores y de multiplicación por escalares.

e. El principio de superposición es una descripción física de una transformación lineal.

23. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal que refleja cada punto a través del eje x_1. (Vea el problema de práctica 2.) Trace dos bosquejos similares a la figura 6 que ilustra las propiedades (i) y (ii) de una transformación lineal.

24. Suponga que los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_p$ generan $\mathbb{R}^n$ y sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal. suponga que $T(\mathbf{v}_i) = \mathbf{0}$ para $i = 1, \ldots, p$. Muestre que T es la transformación cero. Esto es, muestre que si $\mathbf{x}$ es cualquier vector en $\mathbb{R}^n$, entonces $T(\mathbf{x}) = \mathbf{0}$.

25. Dados $\mathbf{v} \neq \mathbf{0}$ y $\mathbf{p}$ en $\mathbb{R}^n$, la línea que pasa por $\mathbf{p}$ en la dirección de $\mathbf{v}$ tiene la ecuación paramétrica $\mathbf{x} = \mathbf{p} + t\mathbf{v}$. Muestre que una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ mapea esta línea sobre otra línea o sobre un único punto (una *línea degenerada*).

26. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores linealmente independientes en $\mathbb{R}^3$, y sea P el plano a través de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{0}$. La ecuación paramétrica de P es $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ (con s, t en $\mathbb{R}$). Muestre que una transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ mapea P sobre un plano que pasa por $\mathbf{0}$, sobre una línea que pasa por $\mathbf{0}$, o únicamente sobre el origen en $\mathbb{R}^3$. ¿Qué característica deben tener $T(\mathbf{u})$ y $T(\mathbf{v})$ para que la imagen del plano P sea un plano?

27. a. Muestre que la línea que pasa por los vectores $\mathbf{p}$ y $\mathbf{q}$ en $\mathbb{R}^n$ puede escribirse en la forma paramétrica $\mathbf{x} = (1 - t)\mathbf{p} + t\mathbf{q}$. (Vea la figura que acompaña a los ejercicios 21 y 22 de la sección 1.5.)

b. El segmento de línea de $\mathbf{p}$ a $\mathbf{q}$ es el conjunto de puntos de la forma $(1 - t)\mathbf{p} + t\mathbf{q}$ para $0 \leq t \leq 1$ (como indica la siguiente figura). Muestre que una transformación lineal T mapea este segmento de línea sobre un segmento de línea o sobre un único punto.

$(t = 1)$ $\mathbf{q}$ — $(1 - t)\mathbf{p} + t\mathbf{q}$ — $\mathbf{x}$ — $(t = 0)$ $\mathbf{p}$

28. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en $\mathbb{R}^n$. Es posible mostrar que el conjunto P de todos los puntos del paralelogramo determinado por $\mathbf{u}$ y $\mathbf{v}$ tiene la forma $a\mathbf{u} + b\mathbf{v}$ para $0 \leq a \leq 1$, $0 \leq b \leq 1$. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Explique por qué la imagen de un punto en P bajo la transformación T yace en el paralelogramo determinado por $T(\mathbf{u})$ y $T(\mathbf{v})$.

29. Defina $f : \mathbb{R} \to \mathbb{R}$ como $f(x) = mx + b$.

a. muestre que f es una transformación lineal cuando $b = 0$.

b. Encuentre una propiedad de una transformación lineal que se viole cuando $b \neq 0$.

c. ¿Por qué se dice que f es una función lineal?

30. Una *transformación afín* $T : \mathbb{R}^n \to \mathbb{R}^m$ tiene la forma $T(x) = A\mathbf{x} + \mathbf{b}$, donde A es una matriz de $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$. Muestre que T *no* es una transformación lineal cuando $\mathbf{b} \neq \mathbf{0}$. (Las transformaciones afines son importantes en la graficación por computadora.)

31. Sean $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal y $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ un conjunto linealmente dependiente en $\mathbb{R}^n$. Explique por qué el conjunto $\{T(\mathbf{v}_1), T(\mathbf{v}_2), T(\mathbf{v}_3)\}$ es linealmente dependiente.

En los ejercicios 32 a 36, los vectores columna se escriben como filas, por ejemplo $\mathbf{x} = (x_1, x_2)$, y $T(\mathbf{x})$ se escribe como $T(x_1, x_2)$.

32. Muestre que la transformación T definida por $T(x_1, x_2) = (4x_1 - 2x_2, 3|x_2|)$ no es lineal.

33. Muestre que la transformación T definida por $T(x_1, x_2) = (2x_1 - 3x_2, x_1 + 4, 5x_2)$ no es lineal.

34. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Muestre que si T mapea dos vectores linealmente independientes sobre un conjunto linealmente dependiente, entonces la ecuación $T(\mathbf{x}) = \mathbf{0}$ tiene una solución no trivial. [*Sugerencia:* Suponga que $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ son linealmente independientes, pero que $T(\mathbf{u})$ y $T(\mathbf{v})$ son linealmente dependientes. Entonces $c_1T(\mathbf{u}) + c_2T(\mathbf{v}) = \mathbf{0}$ para algunos pesos c_1 y c_2, donde al menos uno de ellos no es cero. Use esta ecuación.]

35. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación que refleja cada vector $\mathbf{x} = (x_1, x_2, x_3)$ a través del plano $x_3 = 0$ sobre $T(\mathbf{x}) = (x_1, x_2, -x_3)$. Muestre que T es una transformación lineal. [Para adquirir algunas ideas útiles vea el ejemplo 4.]

36. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación que proyecta cada vector $\mathbf{x} = (x_1, x_2, x_3)$ sobre el plano $x_2 = 0$, de modo que $T(\mathbf{x}) = (x_1, 0, x_3)$. Muestre que T es una transformación lineal.

[M] En los ejercicios 37 y 38, la matriz dada determina una transformación lineal T. Encuentre todas las $\mathbf{x}$ que satisfagan $T(\mathbf{x}) = \mathbf{0}$.

37. $\begin{bmatrix} 4 & -2 & 5 & -5 \\ -9 & 7 & -8 & 0 \\ -6 & 4 & 5 & 3 \\ 5 & -3 & 8 & -4 \end{bmatrix}$

38. $\begin{bmatrix} -9 & -4 & -9 & 4 \\ 5 & -8 & -7 & 6 \\ 7 & 11 & 16 & -9 \\ 9 & -7 & -4 & 5 \end{bmatrix}$

39. **[M]** Sea $\mathbf{b} = \begin{bmatrix} 7 \\ 5 \\ 9 \\ 7 \end{bmatrix}$ y A la matriz del ejercicio 37. ¿Está $\mathbf{b}$ en el rango de la transformación $\mathbf{x} \mapsto A\mathbf{x}$? Si es así, encuentre una $\mathbf{x}$ cuya imagen bajo la transformación sea $\mathbf{b}$.

40. **[M]** Sea $\mathbf{b} = \begin{bmatrix} -7 \\ -7 \\ 13 \\ -5 \end{bmatrix}$ y A la matriz del ejercicio 38. ¿Está $\mathbf{b}$ en el rango de la transformación $\mathbf{x} \mapsto A\mathbf{x}$? Si es así, encuentre una $\mathbf{x}$ cuya imagen bajo la transformación sea $\mathbf{b}$.

SG Dominio de las transformaciones lineales 1 a 37 (Mastering: Linear Transformations 1-37)

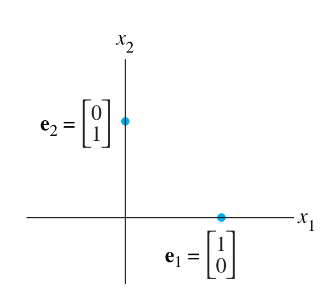
La transformación $\mathbf{x} \mapsto A\mathbf{x}$.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. A debe tener cinco columnas para que $A\mathbf{x}$ esté definida. A debe tener dos filas para que el codominio de T sea $\mathbb{R}^2$.
2. Grafique algunos puntos aleatorios (vectores) en papel para graficar a fin de observar lo que sucede. Un punto como (4, 1) se mapea a (4, − 1). La transformación $\mathbf{x} \mapsto A\mathbf{x}$ refleja puntos a través del eje x (o eje x_1).
3. Sea $\mathbf{x} = t\mathbf{u}$ para alguna t, de tal forma que $0 \leq t \leq 1$. Como T es lineal, $T(t\mathbf{u}) = t\,T(\mathbf{u})$, que es el punto existente sobre el segmento de recta que está entre $\mathbf{0}$ y $T(\mathbf{u})$.

1.9 LA MATRIZ DE UNA TRANSFORMACIÓN LINEAL

Siempre que una transformación lineal T surge de manera geométrica o se describe con palabras, es común desear tener una "fórmula" para $T(\mathbf{x})$. En el análisis siguiente se muestra que toda transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ es en realidad una transformación matricial $\mathbf{x} \mapsto A\mathbf{x}$, y que algunas propiedades importantes de T están íntimamente relacionadas con propiedades conocidas de A. La clave para encontrar A es observar que T está completamente determinada por lo que le hace a las columnas de la matriz identidad $n \times n$, I_n.



EJEMPLO 1 Las columnas de $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ son $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Suponga que T es una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^3$ de tal modo que

$$T(\mathbf{e}_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad \text{y} \quad T(\mathbf{e}_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

Sin más información, encuentre una fórmula para la imagen de una $\mathbf{x}$ arbitraria en $\mathbb{R}^2$.

Solución Escriba

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \tag{1}$$

Como T es una transformación *lineal*,

$$T(\mathbf{x}) = x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) \qquad (2)$$

$$= x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix}$$

El paso de (1) a (2) explica por qué el hecho de conocer $T(\mathbf{e}_1)$ y $T(\mathbf{e}_2)$ es suficiente para poder determinar $T(\mathbf{x})$ para cualquier $\mathbf{x}$. Además, puesto que (2) expresa $T(\mathbf{x})$ como una combinación lineal de vectores, se pueden poner estos vectores en las columnas de una matriz A y escribir (2) como

$$T(\mathbf{x}) = [\, T(\mathbf{e}_1) \quad T(\mathbf{e}_2)\,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A\mathbf{x}$$

TEOREMA 10

Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Entonces existe una única matriz A tal que

$$T(\mathbf{x}) = A\mathbf{x} \text{ para toda } \mathbf{x} \text{ en } \mathbb{R}^n$$

De hecho, A es la matriz de $m \times n$ cuya j-ésima columna es el vector $T(\mathbf{e}_j)$, donde $\mathbf{e}_j$ es la j-ésima columna de la matriz identidad en $\mathbb{R}^n$.

$$A = [T(\mathbf{e}_1) \quad \cdots \quad T(\mathbf{e}_n)] \qquad (3)$$

DEMOSTRACIÓN Escriba $\mathbf{x} = I_n\mathbf{x} = [\mathbf{e}_1 \quad \cdots \quad \mathbf{e}_n]\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n e_n$, y use la linealidad de T para calcular

$$T(\mathbf{x}) = T(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1 T(\mathbf{e}_1) + \cdots + x_n T(\mathbf{e}_n)$$

$$= [\, T(\mathbf{e}_1) \quad \cdots \quad T(\mathbf{e}_n)\,] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A\mathbf{x}$$

La unicidad de A se considera en el ejercicio 33.

La matriz A en (3) se denomina **matriz canónica para la transformación lineal** T.

Ahora se sabe que cada transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ es una transformación matricial, y viceversa. El término *transformación lineal* se centra en una propiedad de una función, mientras que el término *transformación matricial* describe cómo se implementa una transformación de este tipo; lo cual se ilustra en los ejemplos siguientes.

EJEMPLO 2 Encuentre la matriz estándar A para la transformación dilatación $T(\mathbf{x}) = 3\mathbf{x}$, para $\mathbf{x}$ en $\mathbb{R}^2$.

Solución Escriba

$$T(\mathbf{e}_1) = 3\mathbf{e}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \quad \text{y} \quad T(\mathbf{e}_2) = 3\mathbf{e}_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

EJEMPLO 3 Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación que gira cada punto en $\mathbb{R}^2$ un ángulo φ, el cual es positivo si va en dirección contraria al movimiento de las manecillas del reloj. Se podría mostrar geométricamente que dicha transformación es lineal. (Vea la figura 6 de la sección 1.8.) Encuentre la matriz estándar A para esta transformación.

Solución $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ gira a $\begin{bmatrix} \cos\varphi \\ \operatorname{sen}\varphi \end{bmatrix}$, y $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ gira a $\begin{bmatrix} -\operatorname{sen}\varphi \\ \cos\varphi \end{bmatrix}$. Vea la figura 1. De acuerdo con el teorema 10,

$$A = \begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}$$

El ejemplo 5 de la sección 1.8 es un caso especial de esta transformación, con $\varphi = \pi/2$.

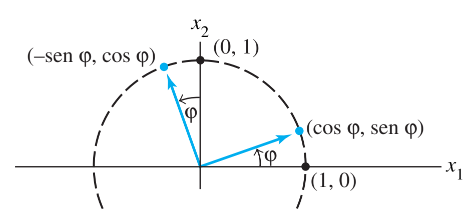


FIGURA 1 Una transformación rotación.

Transformaciones lineales geométricas de $\mathbb{R}^2$

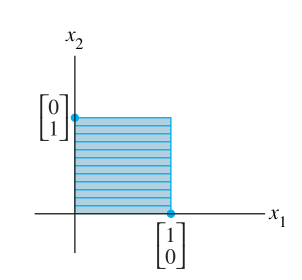


FIGURA 2
El cuadrado unitario.

En los ejemplos 2 y 3 se ilustraron transformaciones lineales que se describen geométricamente. Las tablas 1 a 4 ilustran otras transformaciones lineales geométricas comunes del plano. Debido a que las transformaciones son lineales, quedan completamente determinadas por lo que hacen a las columnas de I_2. En vez de mostrar solamente las imágenes de $\mathbf{e}_1$ y $\mathbf{e}_2$, las tablas incluyen lo que una transformación hace a un cuadrado unitario (figura 2).

Se pueden construir otras transformaciones aparte de las enlistadas en las tablas de la 1 a la 4, siempre y cuando se aplique una transformación después de otra. Por ejemplo, una transformación de trasquilado horizontal puede ir seguida de una reflexión sobre el eje x_2. En la sección 2.1 se mostrará que una *composición* de transformaciones lineales de este tipo es lineal. (Vea también el ejercicio 36.)

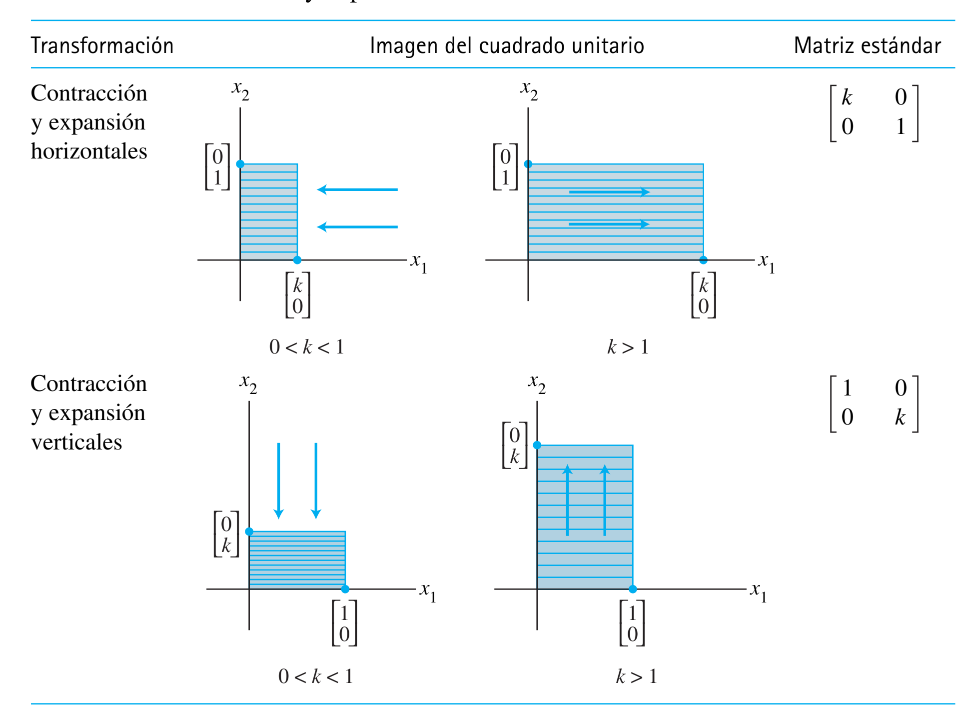
Preguntas de existencia y unicidad

El concepto de transformación lineal ofrece una nueva manera de entender las preguntas de existencia y unicidad planteadas en los inicios de este capítulo. Las dos definiciones que siguen a las tablas 1 a 4 proporcionan la terminología apropiada para referirse a las transformaciones.

TABLA 1 Reflexiones

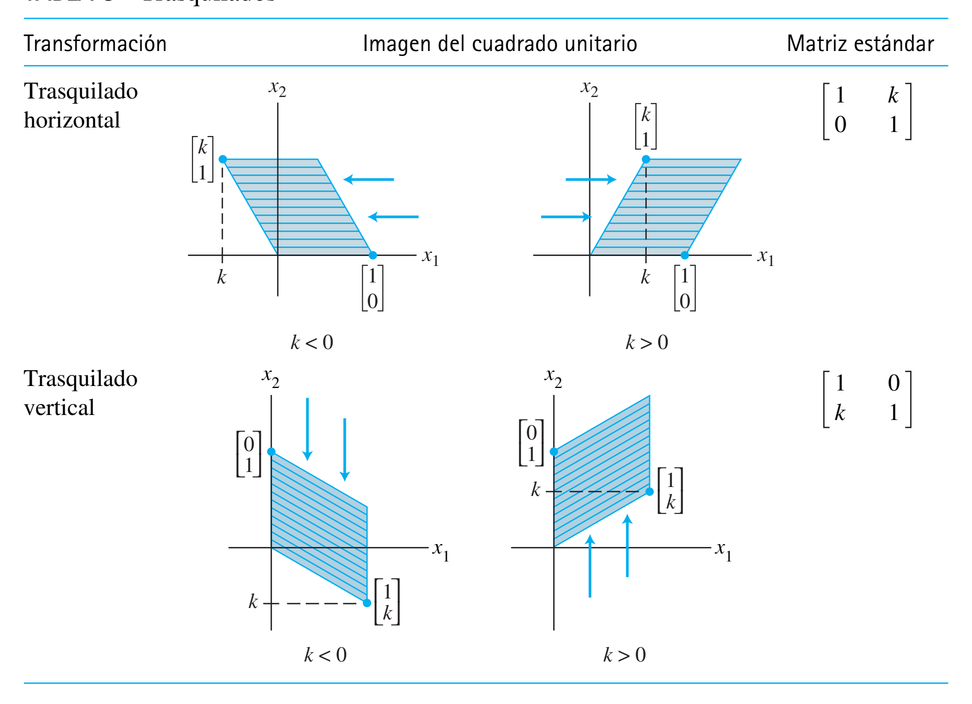
Transformación	Imagen del cuadrado unitario	Matriz estándar
Reflexión a través del eje x_1		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflexión a través del eje x_2		$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflexión a través de la recta $x_2 = x_1$		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Reflexión a través de la recta $x_2 = -x_1$		$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$
Reflexión a través del origen		$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

TABLA 2 Contracciones y expansiones



Transformación	Imagen del cuadrado unitario	Matriz estándar
Contracción y expansión horizontales	$0 < k < 1$; $k > 1$	$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$
Contracción y expansión verticales	$0 < k < 1$; $k > 1$	$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

TABLA 3 Trasquilados



Transformación	Imagen del cuadrado unitario	Matriz estándar
Trasquilado horizontal	$k < 0$; $k > 0$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
Trasquilado vertical	$k < 0$; $k > 0$	$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

TABLA 4 Proyecciones

Transformación	Imagen del cuadrado unitario	Matriz estándar
Proyección sobre el eje x_1	x_2, x_1, $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Proyección sobre el eje x_2	x_2, x_1, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

DEFINICIÓN

Se dice que un mapeo $T : \mathbb{R}^n \to \mathbb{R}^m$ es **sobre** $\mathbb{R}^m$ **(suprayectiva)** si cada **b** en $\mathbb{R}^m$ es la imagen de *al menos una* **x** en $\mathbb{R}^n$.

De manera equivalente, T es sobre $\mathbb{R}^m$ cuando todo el rango de T es todo el codominio $\mathbb{R}^m$. Esto es, T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$ si, para cada **b** en el codominio $\mathbb{R}^m$, existe por lo menos una solución de $T(\mathbf{x}) = \mathbf{b}$. La pregunta "¿mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$?" es una pregunta de existencia. La función T *no* es suprayectiva cuando existe alguna **b** en $\mathbb{R}^m$ tal que la ecuación $T(\mathbf{x}) = \mathbf{b}$ no tenga solución. Vea la figura 3.

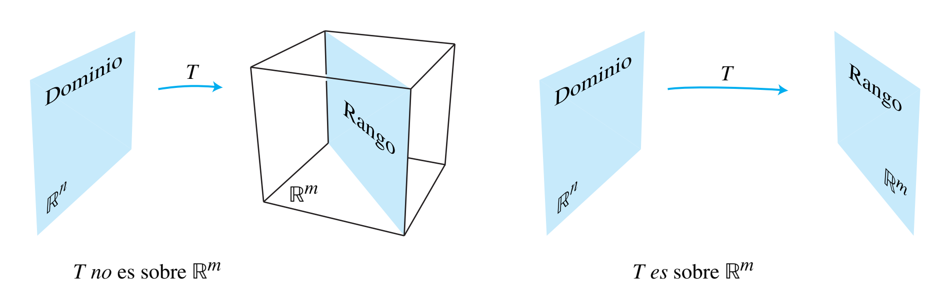


FIGURA 3 ¿El rango de T es todo $\mathbb{R}^m$?

DEFINICIÓN

Una función $T : \mathbb{R}^n \to \mathbb{R}^m$ es **uno a uno (inyectiva)** si cada **b** en $\mathbb{R}^m$ es la imagen de *cuando mucho una* **x** en $\mathbb{R}^n$.

De manera equivalente, T es uno a uno si para cada **b** en $\mathbb{R}^m$ la ecuación $T(\mathbf{x}) = \mathbf{b}$ tiene o una solución única o ninguna solución. La pregunta "¿T es uno a uno?", es una pregunta de unicidad. La función T *no* es uno a uno cuando alguna **b** presente en $\mathbb{R}^m$ es la imagen de más de un vector presente en $\mathbb{R}^n$. Si no existe una **b** con esta característica, entonces T es uno a uno. Vea la figura 4.

SG Dominio de existencia y unicidad 1 a 42 (Mastering: Existence and Uniqueness 1-42)

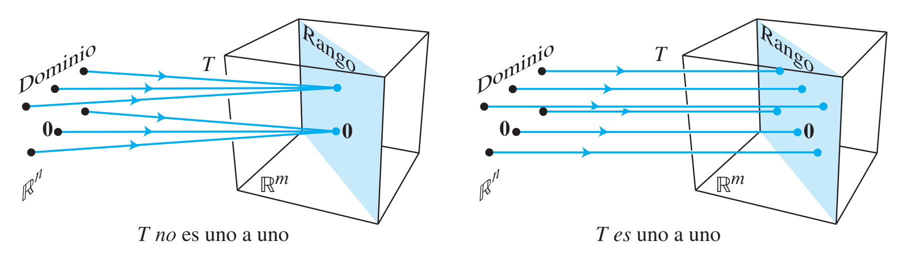


FIGURA 4 ¿Cada **b** es la imagen de, cuando mucho, un vector?

Las transformaciones de proyección mostradas en la tabla 4 *no* son uno a uno y *no* mapean $\mathbb{R}^2$ sobre $\mathbb{R}^2$. Las transformaciones de las tablas 1, 2 y 3 son uno a uno y mapean $\mathbb{R}^2$ sobre $\mathbb{R}^2$. En los dos ejemplos siguientes se muestran otras posibilidades.

En el ejemplo 4, y en los teoremas que siguen, se muestra cómo las propiedades de suprayectividad e inyectividad de las funciones están relacionadas con conceptos que se desarrollaron previamente en este capítulo.

EJEMPLO 4 Sea T la transformación lineal cuya matriz estándar es

$$A = \begin{bmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

¿T mapea $\mathbb{R}^4$ sobre $\mathbb{R}^3$? ¿T *es* una función inyectiva?

Solución Como A está en forma escalonada, puede verse de inmediato que tiene una posición de pivote en cada fila. Por el teorema 4 de la sección 1.4, para cada **b** en $\mathbb{R}^3$ la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente. En otras palabras, la transformación lineal T mapea $\mathbb{R}^4$ (su dominio) sobre $\mathbb{R}^3$. Sin embargo, como la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una variable libre (ya que existen cuatro variables y sólo tres variables básicas), cada **b** es la imagen de más de una **x**. Esto es, T *no* es inyectiva.

TEOREMA 11 Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Entonces T es inyectiva si, y sólo si, la ecuación $T(\mathbf{x}) = \mathbf{0}$ tiene únicamente la solución trivial.

DEMOSTRACIÓN Como T es lineal, entonces $T(\mathbf{0}) = \mathbf{0}$. Si T es inyectiva, entonces la ecuación $T(\mathbf{x}) = \mathbf{0}$ tiene cuando mucho una solución y, por lo tanto, únicamente la solución trivial. Si T no es inyectiva, entonces existe una **b** que es la imagen de al menos dos vectores diferentes en $\mathbb{R}^n$ —por ejemplo, **u** y **v**. Esto es, $T(\mathbf{u}) = \mathbf{b}$ y $T(\mathbf{v}) = \mathbf{b}$. Pero entonces, como T es lineal,

$$T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

El vector $\mathbf{u} - \mathbf{v}$ no es cero, puesto que $\mathbf{u} \neq \mathbf{v}$. Por lo tanto, la ecuación $T(\mathbf{x}) = \mathbf{0}$ tiene más de una solución. Así, las dos condiciones del teorema son verdaderas o bien ambas son falsas. ■

TEOREMA 12 Sean $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal y A la matriz estándar para T. Entonces:

a. T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$ si, y sólo si, las columnas de A generan $\mathbb{R}^m$;

b. T es inyectiva si, y sólo si, las columnas de A son linealmente independientes.

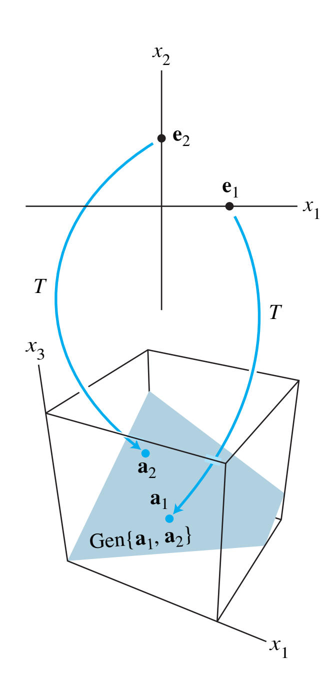
DEMOSTRACIÓN

a. Por el teorema 4 de la sección 1.4, las columnas de A generan $\mathbb{R}^m$ si, y sólo si, para cada $\mathbf{b}$ la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente —en otras palabras, si, y sólo si, para cada $\mathbf{b}$, la ecuación $T(\mathbf{x}) = \mathbf{b}$ tiene por lo menos una solución—. Esto es cierto si T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$.

b. Las ecuaciones $T(\mathbf{x}) = \mathbf{0}$ y $A\mathbf{x} = \mathbf{0}$ son la misma ecuación excepto por la notación. Así, por el teorema 11, T es inyectiva si, y sólo si, $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial. Esto sucede si, y sólo si, las columnas de A son linealmente independientes, como ya se especificó en el enunciado (3) que aparece en un recuadro en la sección 1.7. ■

El enunciado (a) del teorema 12 es equivalente a la afirmación "T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$ si, y sólo si, todo vector en $\mathbb{R}^m$ es una combinación lineal de las columnas de A". Vea el teorema 4 de la sección 1.4.

En el siguiente ejemplo, y en algunos ejercicios subsecuentes, los vectores columna se escriben en filas, como $\mathbf{x} = (x_1, x_2)$, y $T(\mathbf{x})$ se escribe como $T(x_1, x_2)$ en lugar de la manera más formal $T((x_1, x_2))$.



La transformación T no es sobre $\mathbb{R}^3$.

EJEMPLO 5 Sea $T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$. Demuestre que T es una transformación lineal inyectiva. ¿T mapea $\mathbb{R}^2$ sobre $\mathbb{R}^3$?

Solución Cuando $\mathbf{x}$ y $T(\mathbf{x})$ se escriben como vectores columna, la matriz estándar de T puede determinarse por inspección al visualizar el cálculo fila-vector de cada entrada en $A\mathbf{x}$.

$$T(\mathbf{x}) = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad (4)$$

Entonces T es, de hecho, una transformación lineal, y su matriz estándar es la que se muestra en (4). Las columnas de A son linealmente independientes ya que no son múltiplos. Por el teorema 12(b), T es inyectiva. Para decidir si T *es* sobre $\mathbb{R}^3$, examine el espacio generado por las columnas de A. Como A es de 3×2, las columnas de A generan $\mathbb{R}^3$ si, y sólo si, A tiene 3 posiciones pivote, de acuerdo con el teorema 4. Esto es imposible, porque A tiene sólo 2 columnas. Por lo tanto, las columnas de A no generan $\mathbb{R}^3$, y la transformación lineal asociada no es sobre $\mathbb{R}^3$. ■

PROBLEMA DE PRÁCTICA

Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación que primero realiza un trasquilado horizontal que mapea $\mathbf{e}_2$ en $\mathbf{e}_2 - 0.5\mathbf{e}_1$ (pero no modifica $\mathbf{e}_1$) y luego refleja el resultado sobre el eje x_2. Suponiendo que T es lineal, encuentre su matriz estándar. [*Sugerencia:* Determine la localización final de las imágenes de $\mathbf{e}_1$ y $\mathbf{e}_2$.]

1.9 EJERCICIOS

En los ejercicios 1 a 10, suponga que T es una transformación lineal. Encuentre la matriz estándar para T.

1. $T : \mathbb{R}^2 \to \mathbb{R}^4$, $T(\mathbf{e}_1) = (3, 1, 3, 1)$ y $T(\mathbf{e}_2) = (-5, 2, 0, 0)$, donde $\mathbf{e}_1 = (1, 0)$ y $\mathbf{e}_2 = (0, 1)$.

2. $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(\mathbf{e}_1) = (1, 3)$, $T(\mathbf{e}_2) = (4, -7)$, y $T(\mathbf{e}_3) = (-5, 4)$, donde $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ son las columnas de la matriz identidad de 3×3.

3. $T : \mathbb{R}^2 \to \mathbb{R}^2$ gira puntos (alrededor del origen) a través de $3\pi/2$ radianes (en sentido contrario al de las manecillas del reloj).

4. $T : \mathbb{R}^2 \to \mathbb{R}^4$ gira puntos (alrededor del origen) a través de $-\pi/4$ radianes (en el mismo sentido que las manecillas del reloj). [*Sugerencia:* $T(\mathbf{e}_1) = (1/\sqrt{2}, -1/\sqrt{2})$.]

5. $T : \mathbb{R}^2 \to \mathbb{R}^2$ es una transformación de trasquilado vertical que mapea $\mathbf{e}_1$ en $\mathbf{e}_1 - 2\mathbf{e}_2$, pero no modifica al vector $\mathbf{e}_2$.

6. $T : \mathbb{R}^2 \to \mathbb{R}^2$ es una transformación de trasquilado horizontal que mapea $\mathbf{e}_2$ en $\mathbf{e}_2 + 3\mathbf{e}_1$, pero no modifica al vector $\mathbf{e}_1$.

7. $T : \mathbb{R}^2 \to \mathbb{R}^2$ primero gira puntos en el mismo sentido que las manecillas del reloj en un ángulo de $-3\pi/4$ radianes, y luego refleja puntos a través del eje horizontal x_1. [*Sugerencia:* $T(\mathbf{e}_1) = (-1/\sqrt{2}, 1/\sqrt{2})$.]

8. $T : \mathbb{R}^2 \to \mathbb{R}^2$ primero refleja puntos a través del eje horizontal x_1, y luego a través de la recta $x_2 = x_1$.

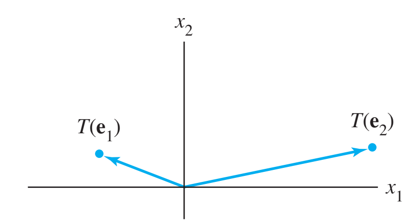
9. $T : \mathbb{R}^2 \to \mathbb{R}^2$ primero realiza un trasquilado horizontal que transforma $\mathbf{e}_2$ en $\mathbf{e}_2 - 2\mathbf{e}_1$ (sin modificar $\mathbf{e}_1$), y luego refleja puntos a través de la recta $x_2 = -x_1$.

10. $T : \mathbb{R}^2 \to \mathbb{R}^2$ primero refleja puntos a través del eje vertical x_2, y luego gira puntos un ángulo de $\pi/2$ radianes.

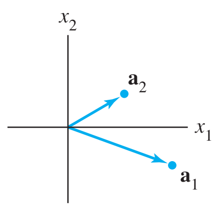
11. Una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ primero refleja puntos a través del eje x_1, y luego a través del eje x_2. Muestre que también T puede describirse como una transformación lineal que gira puntos alrededor del origen. ¿Cuál es el ángulo de ese giro?

12. Muestre que la transformación del ejercicio 8 es en realidad un giro alrededor del origen. ¿Cuál es el ángulo del giro?

13. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal tal que $T(\mathbf{e}_1)$ y $T(\mathbf{e}_2)$ son los vectores mostrados en la figura. Utilice la figura para trazar el vector $T(2, 1)$.



14. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal con matriz estándar $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$, donde $\mathbf{a}_1$ y $\mathbf{a}_2$ se muestran en la figura. Utilice la figura para dibujar la imagen de $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ bajo la transformación T.



En los ejercicios 15 y 16 llene las entradas que faltan en la matriz, suponiendo que las ecuaciones se cumplen para todos los valores de las variables.

15. $$\begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 - 2x_3 \\ 4x_1 \\ x_1 - x_2 + x_3 \end{bmatrix}$$

16. $$\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ -2x_1 + x_2 \\ x_1 \end{bmatrix}$$

En los ejercicios 17 a 20, muestre que T es una transformación lineal encontrando una matriz que implemente la función. Observe que $x_1, x_2, \ldots$ no son vectores sino entradas de vectores.

17. $T(x_1, x_2, x_3, x_4) = (0, x_1 + x_2, x_2 + x_3, x_3 + x_4)$

18. $T(x_1, x_2) = (2x_2 - 3x_1, x_1 - 4x_2, 0, x_2)$

19. $T(x_1, x_2, x_3) = (x_1 - 5x_2 + 4x_3, x_2 - 6x_3)$

20. $T(x_1, x_2, x_3, x_4) = 2x_1 + 3x_3 - 4x_4 \quad (T : \mathbb{R}^4 \to \mathbb{R})$

21. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal tal que $T(x_1, x_2) = (x_1 + x_2, 4x_1 + 5x_2)$. Encuentre una $\mathbf{x}$ tal que $T(\mathbf{x}) = (3, 8)$.

22. Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal tal que $T(x_1, x_2) = (x_1 - 2x_2, -x_1 + 3x_2, 3x_1 - 2x_2)$. Encuentre una $\mathbf{x}$ tal que $T(\mathbf{x}) = (-1, 4, 9)$.

En los ejercicios 23 y 24, señale cada enunciado como verdadero o falso. Justifique cada una de sus respuestas.

23. a. Una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ está completamente determinada por su efecto sobre las columnas de la matriz identidad $n \times n$.

b. Si $T : \mathbb{R}^2 \to \mathbb{R}^2$ gira vectores alrededor del origen en un ángulo φ, entonces T es una transformación lineal.

c. Cuando se realizan dos transformaciones lineales una después de la otra, el efecto combinado puede no ser siempre una transformación lineal.

d. Una función $T : \mathbb{R}^n \to \mathbb{R}^m$ es sobre $\mathbb{R}^m$ si cada vector $\mathbf{x}$ en $\mathbb{R}^n$ se mapea sobre algún vector en $\mathbb{R}^m$.

e. Si A es una matriz de 3×2, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no puede ser uno a uno.

24. a. No toda transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ es una transformación matricial.

b. Las columnas de la matriz estándar para una transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ son las imágenes de las columnas de la matriz identidad $n \times n$.

c. La matriz estándar de una transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^2$ que refleja puntos a través del eje horizontal, el eje vertical o el origen tiene la forma $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$, donde a y d son ± 1.

d. Una función $T : \mathbb{R}^n \to \mathbb{R}^m$ es inyectiva si cada vector en $\mathbb{R}^n$ se mapea sobre un único vector en $\mathbb{R}^m$.

e. Si A es una matriz de 3×2, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no puede mapear $\mathbb{R}^2$ sobre $\mathbb{R}^3$.

En los ejercicios 25 a 28, determine si la transformación lineal es (a) inyectiva, (b) suprayectiva. Justifique cada respuesta.

25. La transformación del ejercicio 17.

26. La transformación del ejercicio 2.

27. La transformación del ejercicio 19.

28. La transformación del ejercicio 14.

En los ejercicios 29 y 30, describa las posibles formas escalonadas de la matriz estándar para una transformación lineal T. Utilice la notación del ejemplo 1 en la sección 1.2.

29. $T : \mathbb{R}^3 \to \mathbb{R}^4$ es inyectiva.

30. $T : \mathbb{R}^4 \to \mathbb{R}^3$ es suprayectiva.

31. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal con matriz estándar A. Complete el siguiente enunciado para hacerlo verdadero: "T es inyectiva si, y sólo si, A tiene ____ columnas pivote". Explique por qué el enunciado es verdadero. [*Sugerencia:* Vea los ejercicios de la sección 1.7.]

32. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal con matriz estándar A. Complete el siguiente enunciado para hacerlo verdadero: "T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$ si, y sólo si, A tiene ______ columnas pivote". Encuentre algunos teoremas que expliquen por qué el enunciado es verdadero.

33. Verifique la unicidad de A en el teorema 10. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal tal que $T(\mathbf{x}) = B\mathbf{x}$ para alguna matriz B de $m \times n$. Demuestre que si A es la matriz estándar para T, entonces $A = B$. [*Sugerencia:* Muestre que A y B tienen las mismas columnas.]

34. ¿Por qué la pregunta sobre si la transformación lineal T es proyectiva es una pregunta de existencia?

35. Si una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ mapea $\mathbb{R}^n$ *sobre* $\mathbb{R}^m$, ¿puede darse alguna relación entre m y n? Si T es inyectiva, ¿qué se puede decir de m y n?

36. Sean $S : \mathbb{R}^p \to \mathbb{R}^n$ y $T : \mathbb{R}^n \to \mathbb{R}^m$ transformaciones lineales. Muestre que la función $\mathbf{x} \mapsto T(S(\mathbf{x}))$ es una transformación lineal (de $\mathbb{R}^p$ a $\mathbb{R}^m$). [*Sugerencia:* Calcule $T(S(c\mathbf{u} + d\mathbf{v}))$ para $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^p$ y los escalares c y d. Justifique cada paso del cálculo, y explique por qué éste conduce a la conclusión deseada.]

[M] En los ejercicios 37 a 40, sea T una transformación lineal cuya matriz estándar está dada. En los ejercicios 37 y 38, decida si T es una función inyectiva. En los ejercicios 39 y 40, decida si T mapea $\mathbb{R}^5$ sobre $\mathbb{R}^5$. Justifique sus respuestas.

37. $\begin{bmatrix} -5 & 10 & -5 & 4 \\ 8 & 3 & -4 & 7 \\ 4 & -9 & 5 & -3 \\ -3 & -2 & 5 & 4 \end{bmatrix}$

38. $\begin{bmatrix} 7 & 5 & 4 & -9 \\ 10 & 6 & 16 & -4 \\ 12 & 8 & 12 & 7 \\ -8 & -6 & -2 & 5 \end{bmatrix}$

39. $\begin{bmatrix} 4 & -7 & 3 & 7 & 5 \\ 6 & -8 & 5 & 12 & -8 \\ -7 & 10 & -8 & -9 & 14 \\ 3 & -5 & 4 & 2 & -6 \\ -5 & 6 & -6 & -7 & 3 \end{bmatrix}$

40. $\begin{bmatrix} 9 & 13 & 5 & 6 & -1 \\ 14 & 15 & -7 & -6 & 4 \\ -8 & -9 & 12 & -5 & -9 \\ -5 & -6 & -8 & 9 & 8 \\ 13 & 14 & 15 & 2 & 11 \end{bmatrix}$

CD Visualización de transformaciones lineales (Visualizing Linear Transformations)

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Observe lo que sucede a $\mathbf{e}_1$ y $\mathbf{e}_2$. Vea la figura 5. Primero, a $\mathbf{e}_1$ no le afecta el trasquilado y luego se refleja en $-\mathbf{e}_1$. Así, $T(\mathbf{e}_1) = -\mathbf{e}_1$. Segundo, $\mathbf{e}_2$ pasa a $\mathbf{e}_2 - .5\mathbf{e}_1$ después de la transformación de trasquilado. Como una reflexión sobre el eje x_2 convierte $\mathbf{e}_1$ en $-\mathbf{e}_1$ y no modifica $\mathbf{e}_2$, el vector $\mathbf{e}_2 - 0.5\mathbf{e}_1$ pasa a $\mathbf{e}_2 + 0.5\mathbf{e}_1$. Así, $T(\mathbf{e}_2) = \mathbf{e}_2 + .5\mathbf{e}_1$. Por lo tanto, la matriz estándar de T es

$$[\,T(\mathbf{e}_1) \quad T(\mathbf{e}_2)\,] = [\,-\mathbf{e}_1 \quad \mathbf{e}_2 + .5\mathbf{e}_1\,] = \begin{bmatrix} -1 & .5 \\ 0 & 1 \end{bmatrix}$$

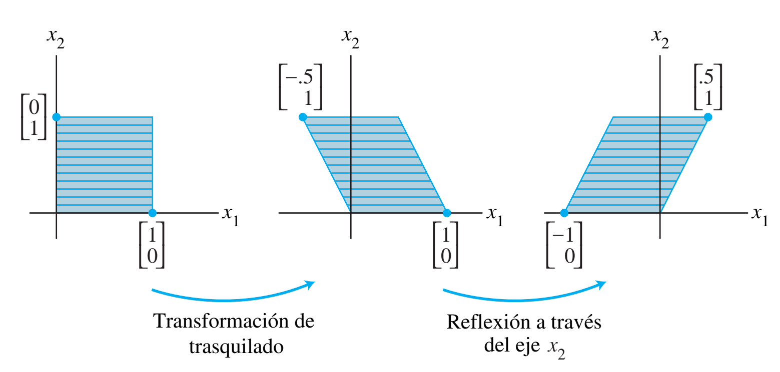


FIGURA 5 La composición de dos transformaciones.

1.10 MODELOS LINEALES EN NEGOCIOS, CIENCIAS E INGENIERÍA

WEB

Todos los modelos matemáticos de esta sección son *lineales*, esto es, cada uno describe un problema por medio de una ecuación lineal, normalmente en forma vectorial o matricial. El primer modelo tiene que ver con nutrición, pero en realidad es representativo de una técnica general para resolver problemas de programación lineal. El segundo modelo proviene de la ingeniería eléctrica. El tercer modelo introduce el concepto de *ecuación lineal en diferencias*, una poderosa herramienta matemática útil para estudiar procesos dinámicos en una amplia variedad de campos como ingeniería, ecología, economía, telecomunicaciones y ciencias administrativas. Los modelos lineales son importantes porque, a menudo, los fenómenos naturales son lineales o casi lineales cuando las variables involucradas se mantienen dentro de fronteras razonables. También, los modelos lineales son más fácilmente adaptables para el cálculo en computadora que los complejos modelos no lineales.

Mientras lea acerca de cada modelo, preste atención a la forma en que su linealidad refleja alguna propiedad del sistema que se modela.

Diseño de una dieta nutritiva para perder peso

La fórmula para la dieta Cambridge, popular en la década de 1980, se basó en años de investigación. Un equipo de científicos, encabezado por el doctor Alan H. Howard, elaboró esta dieta en Cambridge University después de más de ocho años de trabajo clínico con pacientes obesos.[1] La dieta, que consiste en una fórmula en polvo con muy pocas calorías, combina en un equilibrio muy preciso carbohidratos, proteínas de alta calidad y grasa, además de vitaminas, minerales, elementos traza y electrolitos. Millones de personas han usado esta dieta en años recientes para lograr una pérdida de peso rápida y sustancial.

Para encontrar las cantidades y proporciones de nutrimentos deseadas, el doctor Howard tuvo que incorporar una gran variedad de comestibles en la dieta. Cada comestible proporcionaba varios de los ingredientes necesarios, pero no en las proporciones correctas. Por ejemplo, la leche desgrasada era una fuente importante de proteínas, pero contenía demasiado calcio. Por ello se usó harina de soya para conseguir una parte de las proteínas, ya que esta harina contiene muy poco calcio. Sin embargo, la harina de soya aporta una proporción relativamente alta de grasa, así que se agregó suero, pues éste proporciona menos grasa para una cantidad dada de calcio. Desafortunadamente, el suero contiene demasiados carbohidratos. . . .

El ejemplo siguiente ilustra el problema a pequeña escala. En la tabla 1 se mencionan tres de los ingredientes de la dieta, junto con las cantidades de ciertos nutrimentos proporcionadas por 100 gramos de cada ingrediente.[2]

TABLA 1

	Cantidades (en gramos) proporcionadas por 100 g de ingredientes			
Nutrimento	Leche desgrasada	Harina de soya	Suero	Cantidades propocionadas por la dieta Cambridge en un día
Proteínas	36	51	13	33
Carbohidratos	52	34	74	45
Grasa	0	7	1.1	3

EJEMPLO 1 Si es posible, encuentre alguna combinación de leche desgrasada, harina de soya y suero que proporcione las cantidades exactas de proteínas, carbohidratos y grasa proporcionadas por la dieta para un día (tabla 1).

Solución Denote con x_1, x_2 y x_3, respectivamente, los números de unidades (100 gramos) de estos comestibles. Un posible enfoque para encarar el problema es deducir ecuaciones para cada nutrimento por separado. Por ejemplo, el producto

$$\left\{\begin{matrix} x_1 \text{ unidades de} \\ \text{leche desgrasada} \end{matrix}\right\} \cdot \left\{\begin{matrix} \text{proteínas por unidad} \\ \text{de leche desgrasada} \end{matrix}\right\}$$

[1]El primer anuncio de este régimen de pérdida de peso rápida apareció en el *International Journal of Obesity* (1978) **2**, 321–332.

[2]Ingredientes de la dieta en 1984: los datos de nutrimentos de los ingredientes están adaptados de USDA Agricultural Handbooks Núm. 8-1 y 8-6, 1976.

da la cantidad de proteína proporcionada por x_1 unidades de leche desgrasada. A esta cantidad se le agregarían entonces productos similares para harina de soya y suero, y se igualaría la suma resultante a la cantidad de proteínas necesarias. Deben hacerse cálculos análogos para cada nutrimento.

Un método más eficiente, y conceptualmente más simple, es considerar un "vector de nutrimentos" para cada comestible y construir una sola ecuación vectorial. La cantidad de nutrimentos proporcionada por x_1 unidades de leche desgrasada es el múltiplo escalar

$$\overset{\text{Escalar}}{\begin{Bmatrix} x_1 \text{ unidades de} \\ \text{leche desgrasada} \end{Bmatrix}} \cdot \overset{\text{Vector}}{\begin{Bmatrix} \text{nutrimentos por unidad} \\ \text{de leche desgrasada} \end{Bmatrix}} = x_1\mathbf{a}_1 \tag{1}$$

donde $\mathbf{a}_1$ es la primera columna de la tabla 1. Sean $\mathbf{a}_1$ y $\mathbf{a}_3$ los vectores correspondientes para harina de soya y suero, respectivamente, y sea $\mathbf{b}$ el vector que enlista el total de nutrimentos requerido (la última columna de la tabla). Entonces $x_2\mathbf{a}_2$ y $x_3\mathbf{a}_3$ dan los nutrimentos proporcionados por x_2 unidades de harina de soya y x_3 unidades de suero, respectivamente. Así, la ecuación deseada es

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b} \tag{2}$$

La reducción por filas de la matriz aumentada para el sistema de ecuaciones correspondiente muestra que

$$\begin{bmatrix} 36 & 51 & 13 & 33 \\ 52 & 34 & 74 & 45 \\ 0 & 7 & 1.1 & 3 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & .277 \\ 0 & 1 & 0 & .392 \\ 0 & 0 & 1 & .233 \end{bmatrix}$$

Con exactitud de tres dígitos, la dieta requiere de .277 unidades de leche desgrasada, .392 unidades de harina de soya, y .233 unidades de suero para proporcionar las cantidades deseadas de proteínas, carbohidratos y grasa. ■

Es importante que los valores de x_1, x_2 y x_3 encontrados anteriormente sean no negativos. Esto es necesario para que la solución sea factible físicamente. (Por ejemplo, ¿cómo podrían usarse −.233 unidades de suero?) Con un mayor número de requisitos en cuanto a nutrimentos, podría ser necesario usar más cantidades de comestibles para producir un sistema de ecuaciones con una solución "no negativa". Así, podría ser necesario examinar muchísimas combinaciones diferentes de comestibles para encontrar un sistema de ecuaciones con una solución de este tipo. De hecho, el inventor de la dieta Cambridge pudo proporcionar 31 nutrimentos en cantidades precisas usando solamente 33 ingredientes.

El problema de construir una dieta conduce a la ecuación *lineal* (2) porque la cantidad de nutrimentos proporcionada por cada comestible puede escribirse como un múltiplo escalar de un vector, como en (1). Esto es, los nutrimentos aportados por un comestible son *proporcionales* a la cantidad del comestible agregada a la dieta total. También, cada nutrimento de la mezcla es la *suma* de las cantidades de cada comestible.

Los problemas consistentes en formular dietas especializadas para seres humanos y ganado son muy frecuentes. Normalmente se tratan con técnicas de programación lineal. El método para construir ecuaciones vectoriales usado aquí simplifica con frecuencia la tarea de formular esta clase de problemas.

Ecuaciones lineales y redes eléctricas

En una red eléctrica sencilla, la corriente del flujo puede describirse mediante un sistema de ecuaciones lineales. Una fuente de voltaje, por ejemplo una batería, obliga a una corriente de electrones a fluir por la red. Cuando la corriente pasa a través de un resistor (como una bombilla de luz o un motor), una parte del voltaje se "gasta". Según la ley de Ohm, esta "caída de voltaje" a través de un resistor está dada por

$$V = RI$$

donde el voltaje V se mide en *volts*, la resistencia R en *ohms* (denotado por Ω), y el flujo de la corriente I en *amperes.*

La red de la figura 1 contiene tres circuitos cerrados. Las corrientes que fluyen por los circuitos 1, 2 y 3 se denotan mediante I_1, I_2 e I_3, respectivamente. Las direcciones asignadas a tales *corrientes de circuito* son arbitrarias. Si una corriente resulta ser negativa, entonces su dirección real es opuesta a la seleccionada en la figura. Si la dirección de la corriente mostrada es desde el lado positivo (más largo) de una batería (┤⊢) hacia el lado negativo (más corto), el voltaje es positivo; en caso contrario, el voltaje es negativo.

Los flujos de corriente de un circuito están gobernados por la siguiente regla.

> LEY DE KIRCHHOFF DEL VOLTAJE
>
> La suma algebraica de las caídas de voltaje RI en una dirección a lo largo de un circuito es igual a la suma algebraica de las fuentes de voltaje existentes en la misma dirección alrededor del circuito.

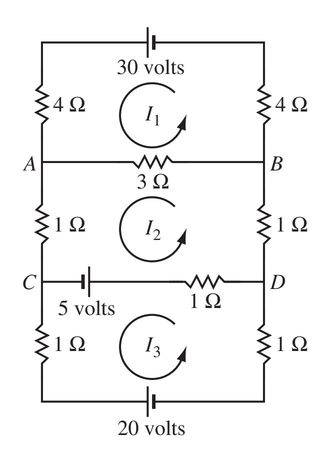


FIGURA 1

EJEMPLO 2 Determine las corrientes de circuito en la red de la figura 1.

Solución Para el circuito 1, la corriente I_1 fluye a través de tres resistores, y la suma de las caídas de voltaje RI es

$$4I_1 + 4I_1 + 3I_1 = (4 + 4 + 3)I_1 = 11I_1$$

La corriente del circuito 2 también fluye en parte del circuito 1, a través de la *rama* corta entre A y B. Aquí, la caída de voltaje RI es de $3I_2$ volts. Sin embargo, la dirección de la corriente para la rama AB del circuito 1 es opuesta a la elegida para el flujo del circuito 2, así que la suma algebraica de todas las caídas RI para el circuito 1 es $11I_1 - 3I_2$. Como el voltaje del circuito 1 es de +30 volts, la ley de Kirchhoff del voltaje implica que

$$11I_1 - 3I_2 = 30$$

La ecuación para el circuito 2 es

$$-3I_1 + 6I_2 - I_3 = 5$$

El término $-3I_1$ proviene del flujo de la corriente del circuito 1 a través de la rama AB (con una caída de voltaje negativa debido a que ahí el flujo de la corriente es opuesto al flujo del circuito 2). El término $6I_2$ es la suma de todas las resistencias del circuito 2, multiplicada por la corriente de circuito. El término $-I_3 = -1 \cdot I_3$ proviene del flujo

de la corriente del circuito 3 a través del resistor de 1 ohm de la rama CD, en dirección opuesta al flujo del circuito 2. La ecuación para el circuito 3 es

$$-I_2 + 3I_3 = -25$$

Observe que la batería de 5 volts de la rama CD se cuenta como parte de ambos circuitos, 2 y 3, pero es de −5 volts para el circuito 3 por la dirección elegida para la corriente del circuito 3. La batería de 20 volts es negativa por la misma razón.

Las corrientes de circuito se encuentran al resolver el sistema

$$\begin{aligned} 11I_1 - 3I_2 \phantom{{}-I_3} &= 30 \\ -3I_1 + 6I_2 - I_3 &= 5 \\ - I_2 + 3I_3 &= -25 \end{aligned} \tag{3}$$

La solución se encuentra al aplicar operaciones por fila: $I_1 = 3$ amperes, $I_2 = 1$ ampere, e $I_3 = -8$ amperes. El valor negativo de I_3 indica que la corriente real en el circuito 3 fluye en dirección opuesta a la indicada en la figura 1.

Resulta instructivo ver el sistema (3) como una ecuación vectorial:

$$I_1 \underset{\mathbf{r}_1}{\underset{\uparrow}{\begin{bmatrix} 11 \\ -3 \\ 0 \end{bmatrix}}} + I_2 \underset{\mathbf{r}_2}{\underset{\uparrow}{\begin{bmatrix} -3 \\ 6 \\ -1 \end{bmatrix}}} + I_3 \underset{\mathbf{r}_3}{\underset{\uparrow}{\begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}}} = \underset{\mathbf{v}}{\underset{\uparrow}{\begin{bmatrix} 30 \\ 5 \\ -25 \end{bmatrix}}} \tag{4}$$

La primera entrada de cada vector está relacionada con el primer circuito, y de manera similar para las entradas segunda y tercera. El primer vector de resistor $\mathbf{r}_1$ enlista la resistencia en los diversos circuitos por los que fluye la corriente I_1. Una resistencia se escribe negativa cuando I_1 fluye contra la dirección de flujo de otro circuito. Examine la figura 1 y observe cómo se calculan las entradas de $\mathbf{r}_1$; después haga lo mismo para $\mathbf{r}_2$ y $\mathbf{r}_3$. La forma matricial de (4),

$$R\mathbf{i} = \mathbf{v}, \quad \text{donde} \quad R = [\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3\,] \quad \text{e} \quad \mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}$$

proporciona una versión matricial de la ley de Ohm. Si todas las corrientes de circuito se eligen en la misma dirección (por ejemplo en dirección contraria al movimiento de las manecillas del reloj), entonces todas las entradas de la diagonal principal de R serán negativas.

La ecuación matricial $R\mathbf{i} = \mathbf{v}$ propicia que la linealidad de este modelo salte a la vista. Por ejemplo, si el vector de voltaje se duplica, el vector de corriente debe ser el doble. También se aplica *un principio de superposición*. Esto es, la solución de la ecuación (4) es la suma de las soluciones de las ecuaciones

$$R\mathbf{i} = \begin{bmatrix} 30 \\ 0 \\ 0 \end{bmatrix}, \qquad R\mathbf{i} = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}, \qquad \text{y} \qquad R\mathbf{i} = \begin{bmatrix} 0 \\ 0 \\ -25 \end{bmatrix}$$

Cada una de estas ecuaciones corresponde al circuito con una sola fuente de voltaje (las otras fuentes se reemplazan por alambres que cierran cada circuito). El modelo para el

flujo de la corriente es *lineal* debido, precisamente, a que las leyes de Ohm y de Kirchhoff son lineales: la caída de voltaje a través de un resistor es *proporcional* al flujo de corriente a través de él (Ohm), y la *suma* de las caídas de voltaje de un circuito es igual a la suma de las fuentes de voltaje del circuito (Kirchhoff).

Las corrientes de circuito presentes en una red pueden servir para determinar la corriente que haya en cualquier rama de la red. Cuando sólo una corriente de circuito pasa por una rama, como de B a D en la figura 1, la corriente presente en la rama es igual a la del circuito. Si más de una corriente de circuito pasa por una rama, como de A a B, la corriente de la rama es la suma algebraica de la corrientes de circuito que haya en la rama (*ley de Kirchhoff de la corriente*). Por ejemplo, la corriente en la rama AB es $I_1 - I_2 = 3 - 1 = 2$ amperes, en la dirección de I_1. La corriente en la rama CD es $I_2 + I_3 = 9$ amperes.

Ecuaciones en diferencias

En muchos campos, tales como ecología, economía e ingeniería, surge la necesidad de modelar matemáticamente un sistema dinámico que cambia a lo largo del tiempo. Algunas características del sistema se miden a intervalos de tiempo discretos, con lo cual se produce una secuencia de vectores $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$. Las entradas de $\mathbf{x}_k$ proporcionan información acerca del *estado* del sistema en el momento de la k-ésima medición.

Si existe una matriz A tal que $\mathbf{x}_1 = A\mathbf{x}_0$, $\mathbf{x}_2 = A\mathbf{x}_1$, y, en general,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{para } k = 0, 1, 2, \ldots \tag{5}$$

entonces (5) se llama **ecuación lineal en diferencias** (o **relación de recurrencia**). Dada una ecuación así, se pueden calcular $\mathbf{x}_1$, $\mathbf{x}_2$, y así sucesivamente, toda vez que $\mathbf{x}_0$ sea conocida. En las secciones 4.8 y 4.9, y en varias secciones del capítulo 5, se desarrollarán fórmulas para $\mathbf{x}_k$ y se describirá qué le sucede a $\mathbf{x}_k$ conforme k se incrementa de manera indefinida. El análisis presentado a continuación ilustra cómo podría surgir una ecuación en diferencias.

Un asunto de interés para los demógrafos es el movimiento de poblaciones o grupos de personas de un lugar a otro. Se considerará aquí un modelo sencillo para los cambios observados en la población de cierta ciudad y sus suburbios durante un periodo de varios años.

Fije un año inicial —por ejemplo 2000— y denote la población de la ciudad y los suburbios mediante r_0 y s_0, respectivamente. Sea $\mathbf{x}_0$ el vector de población

$$\mathbf{x}_0 = \begin{bmatrix} r_0 \\ s_0 \end{bmatrix} \qquad \begin{matrix} \text{Población de la ciudad, 2000} \\ \text{Población de los suburbios, 2000} \end{matrix}$$

Para el 2001 y los años subsecuentes, denote la población de la ciudad y los suburbios mediante los vectores

$$\mathbf{x}_1 = \begin{bmatrix} r_1 \\ s_1 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} r_2 \\ s_2 \end{bmatrix}, \qquad \mathbf{x}_3 = \begin{bmatrix} r_3 \\ s_3 \end{bmatrix}, \ldots$$

El propósito aquí es describir matemáticamente la relación entre estos vectores.

Suponga que los estudios demográficos muestran que, cada año, el 5% de la población de la ciudad se muda a los suburbios (mientras que el 95% permanece en la ciudad), en tanto que el 3% de la población suburbana se muda a la ciudad (y el otro 97% se queda en los suburbios). Vea la figura 2.

Después de un año, la cantidad original r_0 de personas residentes en la ciudad se ha distribuido entre la ciudad y los suburbios de la siguiente manera:

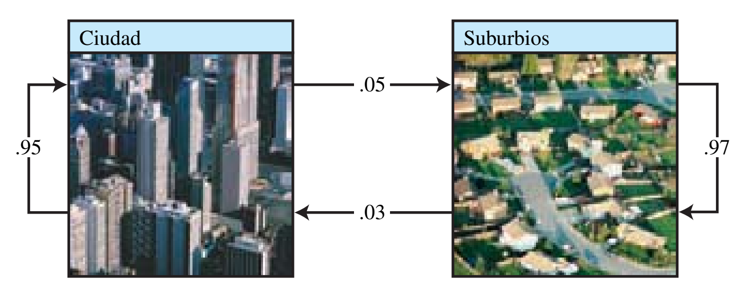


FIGURA 2 Porcentaje anual de migración entre ciudad y suburbios.

$$\begin{bmatrix} .95r_0 \\ .05r_0 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} \quad \begin{matrix} \text{Permanecen en la ciudad} \\ \text{Se mudan a los suburbios} \end{matrix} \tag{6}$$

Las s_0 personas que estaban en los suburbios en el 2000 se distribuyen, después de un año, de la siguiente manera:

$$s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} \quad \begin{matrix} \text{Se mudan a la ciudad} \\ \text{Permanecen en los suburbios} \end{matrix} \tag{7}$$

Los vectores (6) y (7) contabilizan la población total en el 2001.[3] Así que,

$$\begin{bmatrix} r_1 \\ s_1 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} + s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} r_0 \\ s_0 \end{bmatrix}$$

Esto es,

$$\mathbf{x}_1 = M\mathbf{x}_0 \tag{8}$$

donde M es la **matriz de migración** determinada por la tabla siguiente:

$$\begin{matrix} & \text{Desde:} & \\ \text{Ciudad} & \text{Suburbios} & \text{Hacia:} \end{matrix}$$
$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \quad \begin{matrix} \text{Ciudad} \\ \text{Suburbios} \end{matrix}$$

La ecuación (8) describe cómo cambia la población del año 2000 al 2001. Si los porcentajes de migración permanecen constantes, entonces el cambio de 2001 a 2002 está dado por

$$\mathbf{x}_2 = M\mathbf{x}_1$$

y de manera similar para 2002 a 2003 y años subsecuentes. En general,

$$\mathbf{x}_{k+1} = M\mathbf{x}_k \quad \text{para } k = 0, 1, 2, \ldots \tag{9}$$

La secuencia de vectores $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ describe la población existente en la región, ciudad y suburbios, a lo largo de un periodo de años.

---

[3]En aras de la sencillez, se ignoran algunos aspectos que influyen sobre la población, como nacimientos, muertes y emigración e inmigración hacia la región, la cual incluye ciudad y suburbios.

EJEMPLO 3 Determine la población de la región recién descrita para los años 2001 y 2002, si la población en el año 2000 era de 600,000 habitantes en la ciudad y 400,000 en los suburbios.

Solución La población inicial en el año 2000 es $\mathbf{x}_0 = \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$. Para el 2001,

$$\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

Para el 2002,

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix} = \begin{bmatrix} 565{,}440 \\ 434{,}560 \end{bmatrix}$$

El modelo para el movimiento de la población en (9) es *lineal* porque la correspondencia $\mathbf{x}_k \mapsto \mathbf{x}_{k+1}$ es una transformación lineal. La linealidad depende de dos hechos: el número de personas que eligieron mudarse de un área a la otra es *proporcional* al número de personas en el área, tal como se muestra en (6) y en (7), y el efecto acumulado de esas decisiones se obtiene *sumando* los movimientos de las personas desde las diferentes áreas.

PROBLEMAS DE PRÁCTICA

Encuentre una matriz A y vectores $\mathbf{x}$ y $\mathbf{b}$ tales que el problema del ejemplo 1 equivalga a resolver la ecuación $A\mathbf{x} = \mathbf{b}$.

1.10 EJERCICIOS

1. Una caja de cereal para el desayuno indica, normalmente, el número de calorías y las cantidades de proteínas, carbohidratos y grasa contenidas en una porción del cereal. A la derecha se muestran las cantidades para dos conocidos cereales.

Suponga que se debe preparar una mezcla de estos dos cereales que contenga exactamente 295 calorías, 9 g de proteínas, 48 g de carbohidratos, y 8 g de grasa.

a. Establezca una ecuación vectorial para este problema. Incluya un enunciado para explicar qué representa cada variable de la ecuación.

b. Escriba una ecuación matricial equivalente, y luego determine si puede prepararse la mezcla deseada de los dos cereales.

	Información nutricional por porción	
Nutrimento	Cheerios de General Mills	Quaker 100% cereal natural
Calorías	110	130
Proteínas (g)	4	3
Carbohidratos (g)	20	18
Grasa (g)	2	5

2. Una porción (28 g) del salvado de avena Cracklin'Oat Bran proporciona 110 calorías, 3 g de proteínas, 21 g de carbohidratos, y 3 g de grasa. Una porción de Crispix de Kellog's proporciona 110 calorías, 2 g de proteínas, 25 g de carbohidratos, y 0.4 g de grasa.

a. Establezca una matriz B y un vector $\mathbf{u}$ tales que $B\mathbf{u}$ proporcione las cantidades de calorías, proteínas, carbohidratos y grasa contenidas en una mezcla de tres porciones de Cracklin'Oat Bran y dos porciones de Crispix.

b. **[M]** Suponga que se requiere un cereal con más proteínas que Crispix pero menos grasa que Cracklin'Oat Bran. ¿Es posible mezclar los dos cereales para proporcionar 110 calorías, 2.25 g de proteínas, 24 g de carbohidratos, y 1 g de grasa? Si la respuesta es positiva, ¿cuál sería la mezcla?

3. La dieta Cambridge proporciona 0.8 g de calcio por día, además de los nutrimentos enlistados en la tabla 1. Las cantidades de calcio que proporciona una unidad (100 g) de los tres ingredientes de la dieta de Cambridge son: 1.26 g por leche desgrasada, .19 g por harina de soya, y .8 g por suero. Otro ingrediente de la dieta es proteína de soya aislada, la cual proporciona los siguientes nutrimentos por unidad: 80 g de proteínas, 0 g de carbohidratos, 3.4 g de grasa, y .18 g de calcio.

a. Establezca una ecuación matricial cuya solución determine las cantidades de leche desgrasada, harina de soya, suero y proteína de soya aislada necesarias para proporcionar las cantidades exactas de proteínas, carbohidratos, grasa y calcio de la dieta Cambridge. Explique qué representan las variables de la ecuación.

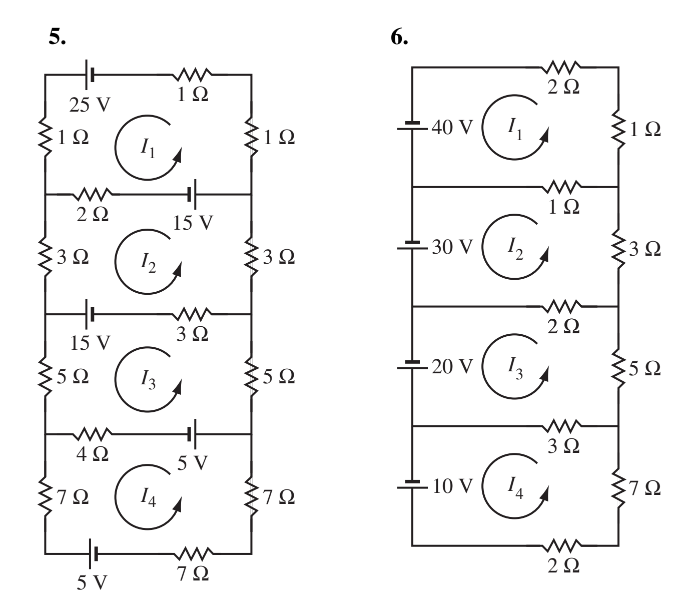
b. **[M]** Resuelva la ecuación de (a) y analice su respuesta.

4. Un dietista está planeando una comida que proporcione ciertas cantidades de vitamina C, calcio y magnesio. Usará tres comestibles y las cantidades se medirán en las unidades apropiadas. Los nutrimentos proporcionados por estos comestibles y los requisitos dietéticos son los siguientes:

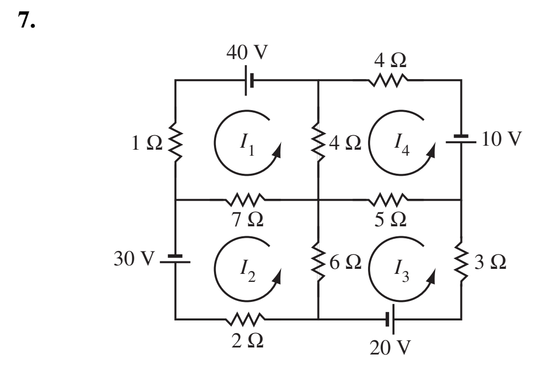
	Miligramos (mg) de nutrimento por unidad de comesitible			Total de nutrientes requeridos (mg)
Nutrimento	Comestible 1	Comestible 2	Comestible 3	
Vitamia C	10	20	20	100
Calcio	50	40	10	300
Magnesio	30	10	40	200

Escriba una ecuación vectorial para este problema. Indique lo que representan las variables y luego resuelva la ecuación.

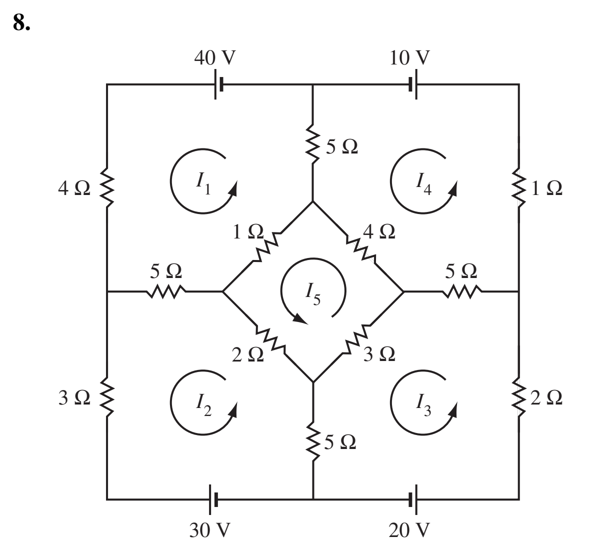
En los ejercicios 5 a 8, escriba una ecuación matricial que determine las corrientes de circuito. **[M]** Si cuenta con MATLAB u otro programa para matrices, resuelva el sistema para las corrientes de circuito.



7.



8.



9. En cierta región, aproximadamente el 5% de la población de la ciudad se muda a los suburbios cada año y cerca del 4% de la población suburbana se muda a la ciudad. En el año 2000 había 600,000 residentes en la ciudad y 400,000 en los suburbios. Establezca una ecuación en diferencias que describa esta situación, donde $\mathbf{x}_0$ sea la población inicial en el 2000. Luego estime la población de la ciudad y de los suburbios dos años después, en 2002. (Ignore otros factores que pudieran influir en los tamaños de las poblaciones.)

10. En cierta región, alrededor del 7% de la población de la ciudad se muda a los suburbios cada año y cerca del 3% de la población suburbana se traslada a la ciudad. En el año 2000 había 800,000 residentes en la ciudad y 500,000 en los suburbios. Establezca una ecuación en diferencias que describa esta situación, donde $\mathbf{x}_0$ sea la población inicial en el 2000. Luego estime la población de la ciudad y de los suburbios dos años después, en 2002.

11. Al iniciar 1990, la población de California era de 29,716,000 habitantes y la de Estados Unidos *fuera* de California era de 218,994,000 habitantes. Durante el año, 509,500 personas se mudaron de California a otra parte de Estados Unidos, mientras que 564,100 personas se mudaron a California desde algún otro lugar de Estados Unidos.[4]

a. Establezca la matriz de migración para esta situación, usando 5 cifras decimales significativas para la migración hacia y desde California. Su trabajo debe mostrar la manera en que se produjo la matriz de migración.

b. **[M]** Determine la población proyectada para el año 2000 de California y del resto de Estados Unidos, suponiendo que las tasas de migración no cambiaron durante el periodo de 10 años. (Estos cálculos no toman en cuenta los nacimientos, las muertes o la importante migración de personas hacia California y otros estados desde fuera de Estados Unidos.)

12. **[M]** La compañía de renta de automóviles Budget con sede en Wichita, Kansas, tiene una flotilla de aproximadamente 450 automóviles en tres sucursales. Un automóvil alquilado en una sucursal puede devolverse en cualquiera de los tres locales. En la tabla que sigue se muestran las distintas proporciones de automóviles devueltos en cada sucursal. Suponga que un lunes hay 304 automóviles en el aeropuerto (o alquilados desde allí), 48 automóviles en la oficina del Este y 98 en la del Oeste. ¿Cuál sería, aproximadamente, la distribución de automóviles el miércoles?

Automóviles alquilados en:

Aeropuerto	Este	Oeste	Devuelto a:
.97	.05	.10	Aeropuerto
.00	.90	.05	Este
.03	.05	.85	Oeste

13. **[M]** Sean M y $\mathbf{x}_0$ como en el ejemplo 3.

a. Determine los vectores de población $\mathbf{x}_k$ para $k = 1, \ldots, 20$. Analice los resultados.

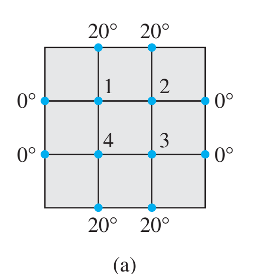
b. Repita (a) con una población inicial de 350,000 habitantes en la ciudad y 650,000 en los suburbios. ¿Qué obtiene?

14. **[M]** Estudie cómo los cambios en las temperaturas de frontera sobre una placa de acero afectan las temperaturas en los puntos interiores de la placa.

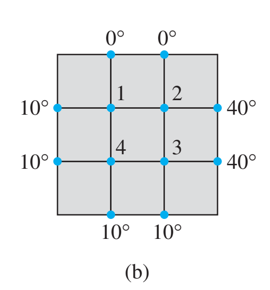
a. Comience por estimar las temperaturas T_1, T_2, T_3, T_4 en cada uno de los conjuntos de cuatro puntos de la placa de acero que se muestra en la figura. En cada caso, el valor de T_k puede aproximarse con el promedio de las temperaturas de los cuatro puntos más cercanos. Vea los ejercicios 33 y 34 de la sección 1.1, donde los valores (en grados) resultan ser (20, 27.5, 30, 22.5). ¿Cómo se relaciona esta lista con los resultados obtenidos para los puntos señalados en el conjunto (a) y el conjunto (b)?

b. Sin realizar ningún cálculo, estime las temperaturas interiores en (a) si todas las temperaturas de frontera se multiplican por 3. Verifique su estimación.

c. Finalmente, formule una conjetura general acerca de la correspondencia de la lista de ocho temperaturas de frontera con la lista de cuatro temperaturas interiores.



(a)



(b)

[4]Datos de migración proporcionados por la Demographic Research Unit del California State Department of Finance.

Solución al problema de práctica

$$A = \begin{bmatrix} 36 & 51 & 13 \\ 52 & 34 & 74 \\ 0 & 7 & 1.1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 33 \\ 45 \\ 3 \end{bmatrix}$$

CAPÍTULO 1 EJERCICIOS SUPLEMENTARIOS

1. Señale cada enunciado como verdadero o falso. Justifique sus respuestas. Si el enunciado es verdadero, cite los hechos o teoremas adecuados; si es falso, explique por qué o dé un contraejemplo de dicha falsedad.

a. Cada matriz es equivalente por filas a una única matriz en forma escalonada.

b. Cualquier sistema de n ecuaciones lineales con n variables tiene cuando mucho n soluciones.

c. Si un sistema de ecuaciones lineales tiene dos soluciones diferentes, debe tener infinidad de soluciones.

d. Si un sistema de ecuaciones lineales no tiene variables libres, entonces tiene una solución única.

e. Si una matriz aumentada $[A \quad \mathbf{b}]$ se transforma en $[C \quad \mathbf{d}]$ mediante operaciones elementales de fila, entonces las ecuaciones $A\mathbf{x} = \mathbf{b}$ y $C\mathbf{x} = \mathbf{d}$ tienen exactamente los mismos conjuntos solución.

f. Si un sistema $A\mathbf{x} = \mathbf{b}$ tiene más de una solución, sucede lo mismo con $A\mathbf{x} = \mathbf{0}$.

g. Si A es una matriz $m \times n$ y la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para alguna $\mathbf{b}$, entonces las columnas de A generan $\mathbb{R}^m$.

h. Si una matriz aumentada $[A \quad \mathbf{b}]$ puede transformarse mediante operaciones elementales de fila a una forma escalonada reducida, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente.

i. Si las matrices A y B son equivalentes por filas, tienen la misma forma escalonada reducida.

j. La ecuación $A\mathbf{x} = \mathbf{0}$ tiene la solución trivial si, y sólo si, no existen variables libres.

k. Si A es una matriz $m \times n$ y la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para toda $\mathbf{b}$ en $\mathbb{R}^m$, entonces A debe tener m columnas pivote.

l. Si una matriz A de $m \times n$ tiene una posición pivote en cada renglón, entonces la ecuación $A\mathbf{x}$ tiene una solución única para cada $\mathbf{b}$ en $\mathbb{R}^m$.

m. Si una matriz A de $n \times n$ tiene n posiciones pivote, entonces la forma escalonada reducida de A es la matriz identidad de $n \times n$.

n. Si las matrices A y B de 3×3 tienen cada una tres posiciones pivote, entonces A puede transformarse en B mediante operaciones elementales de fila.

o. Si A es una matriz $m \times n$, si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos dos soluciones diferentes, y si la ecuación $A\mathbf{x} = \mathbf{c}$ es consistente, entonces la ecuación $A\mathbf{x} = \mathbf{c}$ tiene muchas soluciones.

p. Si A y B son matrices $m \times n$ equivalentes por filas, y si las columnas de A generan $\mathbb{R}^m$, entonces las columnas de B también lo hacen.

q. Si ninguno de los tres vectores en el conjunto $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ de $\mathbb{R}^3$ es un múltiplo de alguno de los otros vectores, entonces S es linealmente independiente.

r. Si $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente independiente, entonces $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ no están en $\mathbb{R}^2$.

s. En algunos casos, es posible que cuatro vectores generen $\mathbb{R}^5$.

t. Si $\mathbf{u}$ y $\mathbf{v}$ están en $\mathbb{R}^m$, entonces $-\mathbf{u}$ está en Gen$\{\mathbf{u}, \mathbf{v}\}$.

u. Si $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ son vectores distintos de cero en $\mathbb{R}^2$, entonces $\mathbf{w}$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$.

v. Si $\mathbf{w}$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, entonces $\mathbf{u}$ es una combinación lineal de $\mathbf{v}$ y $\mathbf{w}$.

w. Suponga que $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ están en $\mathbb{R}^5$, $\mathbf{v}_2$ no es múltiplo de $\mathbf{v}_1$, y $\mathbf{v}_3$ no es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$. Entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente.

x. Una transformación lineal es una función.

y. Si A es una matriz de 6×5, la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ no puede mapear $\mathbb{R}^5$ en $\mathbb{R}^6$.

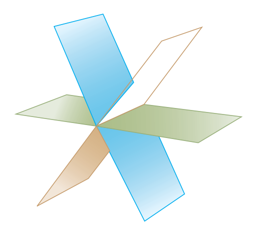
z. Si A es una matriz $m \times n$ con m columnas pivote, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es un mapeo uno a uno.

2. Sean a y b dos números reales. Describa los posibles conjuntos solución de la ecuación (lineal) $ax = b$. [*Sugerencia:* El número de soluciones depende de a y de b.]

3. Las soluciones (x, y, z) de una sola ecuación lineal

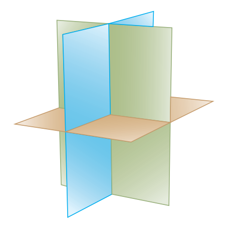
$$ax + by + cz = d$$

forman un plano en $\mathbb{R}^3$ cuando a, b y c no son todas iguales a cero. Construya conjuntos de tres ecuaciones lineales cuyas gráficas (a) se intersecan en una sola línea, (b) se intersecan en un solo punto, (c) no tienen puntos en común. En la figura se ilustran las gráficas típicas.



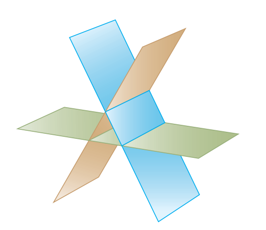
Tres planos que se intersecan en una línea

(a)



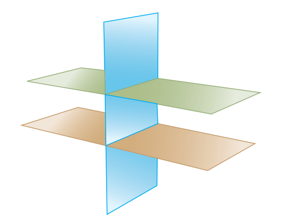
Tres planos que se intersecan en un punto

(b)



Tres planos sin intersección

(c)



Tres planos sin intersección

(c')

4. Suponga que la matriz de coeficientes de un sistema lineal de tres ecuaciones con tres variables tiene un pivote en cada columna. Explique por qué el sistema tiene una solución única.

5. Determine h y k de tal manera que el conjunto solución del sistema (i) sea vacío, (ii) contenga una solución única, y (iii) contenga una infinidad de soluciones.

a. $\begin{aligned} x_1 + 3x_2 &= k \\ 4x_1 + hx_2 &= 8 \end{aligned}$

b. $\begin{aligned} -2x_1 + hx_2 &= 1 \\ 6x_1 + kx_2 &= -2 \end{aligned}$

6. Considere el problema de determinar si el sistema dado a continuación es consistente o no:

$$\begin{aligned} 4x_1 - 2x_2 + 7x_3 &= -5 \\ 8x_1 - 3x_2 + 10x_3 &= -3 \end{aligned}$$

a. Defina vectores apropiados, y replantee el problema en términos de combinaciones lineales. Después, resuelva el problema.

b. Defina una matriz apropiada, y replantee el problema usando la frase "columnas de A".

c. Defina una transformación lineal apropiada T usando la matriz de (b), y replantee el problema en términos de T.

7. Considere el problema de determinar si el siguiente sistema de ecuaciones es consistente para todas b_1, b_2, b_3:

$$\begin{aligned} 2x_1 - 4x_2 - 2x_3 &= b_1 \\ -5x_1 + x_2 + x_3 &= b_2 \\ 7x_1 - 5x_2 - 3x_3 &= b_3 \end{aligned}$$

a. Defina vectores adecuados, y replantee el problema en términos de Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Después, resuelva el problema.

b. Defina una matriz adecuada, y replantee el problema usando la frase "columnas de A".

c. Defina una transformación lineal apropiada T usando la matriz de (b), y replantee el problema en términos de T.

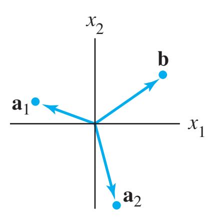
8. Describa las posibles formas escalonadas de la matriz A. Utilice la notación del ejemplo 1 dada en la sección 1.2.

a. A es una matriz de 2×3 cuyas columnas generan $\mathbb{R}^2$.

b. A es una matriz de 3×3 cuyas columnas generan $\mathbb{R}^3$.

9. Escriba el vector $\begin{bmatrix} 5 \\ 6 \end{bmatrix}$ como la suma de dos vectores, uno sobre la línea $\{(x, y): y = 2x\}$, y otro sobre la línea $\{(x, y): y = x/2\}$.

10. Sean $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{b}$ los vectores en $\mathbb{R}^2$ mostrados en la figura, y sea $A = [\mathbf{a}_1 \ \ \mathbf{a}_2]$. ¿La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución? Si es así, ¿la solución es única? Explique su respuesta.



11. Construya una matriz A de 2×3, en forma no escalonada, tal que la solución de $A\mathbf{x} = \mathbf{0}$ sea una línea en $\mathbb{R}^3$.

12. Construya una matriz A de 2×3, en forma no escalonada, tal que la solución de $A\mathbf{x} = \mathbf{0}$ sea un plano en $\mathbb{R}^3$.

13. Escriba la forma escalonada *reducida* de una matriz A de 3×3 tal que las primeras dos columnas de A sean las columnas pivote y $A\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

14. Determine el valor o los valores de a tales que $\left\{\begin{bmatrix} 1 \\ a \end{bmatrix}, \begin{bmatrix} a \\ a+2 \end{bmatrix}\right\}$ sea linealmente independiente.

15. En (a) y en (b), suponga que los vectores son linealmente independientes. ¿Qué puede decir acerca de los números $a, \ldots, f$? Justifique sus respuestas. [*Sugerencia:* Utilice un teorema para (b).]

a. $\begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}, \begin{bmatrix} d \\ e \\ f \end{bmatrix}$

b. $\begin{bmatrix} a \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ c \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} d \\ e \\ f \\ 1 \end{bmatrix}$

16. Use el teorema 7 dado en la sección 1.7 para explicar por qué las columnas de la matriz A son linealmente independientes.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 10 \end{bmatrix}$$

17. Explique por qué un conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ en $\mathbb{R}^5$ debe ser linealmente independiente cuando $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente y $\mathbf{v}_4$ *no* está en Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

18. Suponga que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto linealmente independiente en $\mathbb{R}^n$. Muestre que $\{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2\}$ también es linealmente independiente.

19. Suponga que $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ son puntos distintos de una línea en $\mathbb{R}^3$. No se pide que la línea pase por el origen. Muestre que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente.

20. Sea $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ una transformación lineal, y suponga que $T(\mathbf{u}) = \mathbf{v}$. Muestre que $T(-\mathbf{u}) = -\mathbf{v}$.

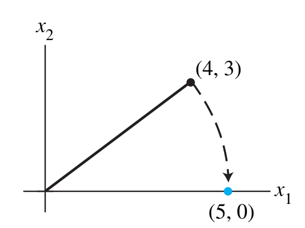
21. Sea $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$ la transformación lineal que refleja cada vector en el plano $x_2 = 0$. Esto es, $T(x_1, x_2, x_3) = (x_1, -x_2, x_3)$. Encuentre la matriz estándar de T.

22. Sea A una matriz de 3×3 con la propiedad de que la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^3$ sobre $\mathbb{R}^3$. Explique por qué la transformación debe ser uno a uno.

23. Una *rotación de Givens* es una transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^n$ que se usa en programas de computadora para crear una entrada cero en un vector (usualmente una columna de una matriz). La matriz estándar de una rotación de Givens en $\mathbb{R}^2$ tiene la forma

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \qquad a^2 + b^2 = 1$$

Encuentre a y b tales que $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ se gira en $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$.



Una rotación de Givens en $\mathbb{R}^2$.

24. La siguiente ecuación describe una rotación de Givens en $\mathbb{R}^3$. Encuentre a y b.

$$\begin{bmatrix} a & 0 & -b \\ 0 & 1 & 0 \\ b & 0 & a \end{bmatrix}\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2\sqrt{5} \\ 3 \\ 0 \end{bmatrix}, \qquad a^2 + b^2 = 1$$

25. Se va a construir un gran edificio de departamentos usando técnicas de construcción modular. La distribución de los departamentos en cualquier piso dado se elige entre tres diseños de piso básicos. El diseño A tiene 18 departamentos en un piso e incluye 3 unidades con tres dormitorios, 7 con dos dormitorios, y 8 con un dormitorio. Cada piso del diseño B incluye 4 unidades con tres dormitorios, 4 con 2 dormitorios, y 8 con un dormitorio. Cada piso del diseño C incluye 5 unidades con tres dormitorios, 3 con dos dormitorios, y 9 con un dormitorio. Suponga que el edificio contiene un total de x_1 pisos del diseño A, x_2 pisos del diseño B, y x_3 pisos del diseño C.

a. ¿Qué interpretación se le puede dar al vector $x_1\begin{bmatrix} 3 \\ 7 \\ 8 \end{bmatrix}$?

b. Escriba una combinación lineal formal de vectores que exprese el número total de departamentos con uno, dos y tres dormitorios existentes en el edificio.

c. **[M]** ¿Es posible diseñar el edificio de tal forma que tenga exactamente 66 unidades con tres dormitorios, 74 unidades con dos dormitorios, y 136 unidades con un dormitorio? Si la respuesta es afirmativa, ¿hay más de una manera de hacerlo? Explique su respuesta.

CD Introducción a MATLAB®

CD Errores de redondeo y pivoteo parcial

CD Introducción al álgebra lineal con Maple®

CD Introducción a Mathematica®

CD Instrucciones básicas para las calculadoras T1-83+, TI-86 y TI-89

CD Introducción a la calculadora HP-48G

2

Álgebra de matrices

EJEMPLO INTRODUCTORIO

Modelos de computadora en el diseño de aviones

Para diseñar la siguiente generación de aviones comerciales y militares, los ingenieros de Phantom Works de Boeing usan el modelado en tres dimensiones y la dinámica de fluidos basada en computadora (CFD, del inglés computational fluid dynamics). Estos profesionales estudian cómo se desplaza el flujo de aire alrededor de un avión virtual para dar respuesta a importantes preguntas sobre el diseño antes de crear modelos físicos. El procedimiento ha reducido en forma drástica los tiempos y costos del ciclo de diseño —y el álgebra lineal desempeña un papel de gran importancia en el proceso.

El avión virtual comienza como un modelo "de alambre" matemático que existe sólo en la memoria de la computadora y en las terminales de despliegue gráfico. (En la ilustración se muestra el modelo de un Boeing 777.) Este modelo matemático organiza e influye en cada paso del diseño y la fabricación del avión —tanto en el exterior como en el interior—. El análisis de CFD tiene que ver con la superficie externa.

Aunque el acabado del forro de un avión puede parecer suave, la geometría de la superficie es complicada. Además de alas y fuselaje, un avión tiene barquillas, estabilizadores, tablillas, aletas y alerones. La forma en que el aire fluye alrededor de estas estructuras determina cómo se mueve el avión en el cielo. Las ecuaciones que describen el flujo del aire son complicadas, y deben tomar en cuenta la admisión de los motores, los gases despedidos por éstos, y las estelas que dejan las alas del avión. Para estudiar el flujo del aire, los ingenieros necesitan de una descripción muy depurada de la superficie del avión.

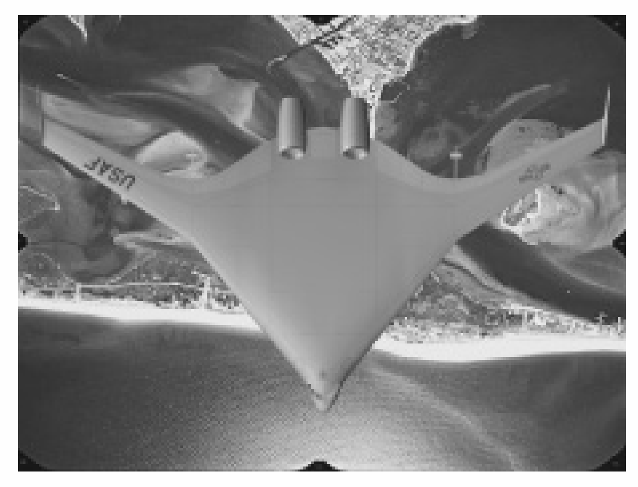
Una computadora crea un modelo de la superficie al superponer, primero, una malla tridimensional de "cuadros" sobre el modelo de alambre original. En esta malla, los cuadros caen completamente dentro o completamente fuera del avión, o intersecan la superficie del mismo. La computadora selecciona los cuadros que intersecan la superficie y los subdivide, reteniendo sólo aquellos más pequeños que aún intersecan la superficie. El proceso de subdivisión se repite hasta que la malla se vuelve extremadamente fina. Una malla típica puede incluir más de 400,000 cuadros.

El proceso para encontrar el flujo de aire alrededor del avión implica la resolución repetida de un sistema de

ecuaciones lineales $A\mathbf{x} = \mathbf{b}$ que puede involucrar hasta 2 millones de ecuaciones y variables. El vector $\mathbf{b}$ cambia a cada momento, con base en datos provenientes de la malla y de las soluciones de ecuaciones previas. Con el uso de computadoras comerciales más rápidas, un equipo de Phantom Works puede emplear desde unas cuantas horas hasta varios días para configurar y resolver un solo problema de flujo de aire. Después, el equipo analiza la solución, puede hacer pequeños cambios a la superficie del avión, y comienza de nuevo con todo el proceso. Pueden requerirse miles de corridas de CFD.

En este capítulo se presentan dos conceptos importantes que ayudan en la resolución de los enormes sistemas de ecuaciones de este tipo:

- *Matrices partidas:* Un sistema de ecuaciones típico de CFD tiene una matriz de coeficientes "dispersa o rala" con entradas que en su mayoría son iguales a cero. El agrupamiento correcto de las variables conduce a una matriz partida con muchos bloques de ceros. En la sección 2.4 se introducen este tipo de matrices y se describen algunas de sus aplicaciones.
- *Factorizaciones de matrices:* Aunque el sistema esté escrito con matrices partidas, sigue siendo complicado. Para simplificar aún más los cálculos, el programa computacional de CFD aplicado en Boeing utiliza lo que se conoce como factorización LU de la matriz de coeficientes. En la sección 2.5 se analiza la factorización LU y otros útiles procedimientos similares. Más adelante, en diversos puntos de este texto, aparecen otros detalles referentes a las factorizaciones.



En la actualidad, la CFD ha revolucionado el diseño de alas. El Boeing Blended Wing Body se encuentra en diseño para ser producido a más tardar en el año 2020.

Para analizar la solución de un sistema de flujo de aire, los ingenieros desean visualizar cómo fluye el aire sobre la superficie del avión. Los ingenieros utilizan gráficas, y el álgebra lineal proporciona el método para elaborarlas. El modelo de alambre de la superficie del avión se almacena como datos en muchas matrices. Una vez que la imagen se despliega en una pantalla de computadora, los ingenieros pueden escalarla, acercar y alejar regiones pequeñas, y girarla para ver partes que pudieran quedar ocultas en determinado ángulo. Cada una de estas operaciones se realiza mediante una multiplicación de matrices adecuada. En la sección 2.7 se explican las ideas básicas de este proceso.

La capacidad para analizar y resolver ecuaciones aumentará considerablemente cuando se adquiera la habilidad de realizar operaciones algebraicas con matrices. Más aún, las definiciones y teoremas de este capítulo proporcionan algunas herramientas básicas para manejar las múltiples aplicaciones del álgebra lineal que involucran a dos o más matrices. Para las matrices cuadradas, el teorema de la matriz invertible presentado en la sección 2.3 reúne la mayor parte de los conceptos tratados an-

teriormente en este texto. En las secciones 2.4 y 2.5 se examinan las matrices partidas y las factorizaciones de matrices que aparecen en la mayor parte de los usos modernos del álgebra lineal. En las secciones 2.6 y 2.7 se describen dos aplicaciones interesantes del álgebra matricial: a la economía y a los gráficos por computadora.

2.1 OPERACIONES DE MATRICES

Si A es una matriz $m \times$ n, esto es, una matriz con m filas y n columnas, entonces la entrada escalar en la i-ésima fila y la j-ésima columna de A se denota mediante a_{ij} y se llama entrada (i, j) de A. Vea la figura 1. Por ejemplo, la entrada (3, 2) es el número a_{32} en la tercera fila, segunda columna. Las columnas de A son vectores en $\mathbb{R}^m$ y se denotan mediante $\mathbf{a}_1, \ldots, \mathbf{a}_n$ (en letras negritas). La atención se centra sobre estas columnas cuando se escribe

$$A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n\,]$$

Observe que el número a_{ij} es la i-ésima entrada (de arriba a abajo) del j-ésimo vector columna $\mathbf{a}_j$.

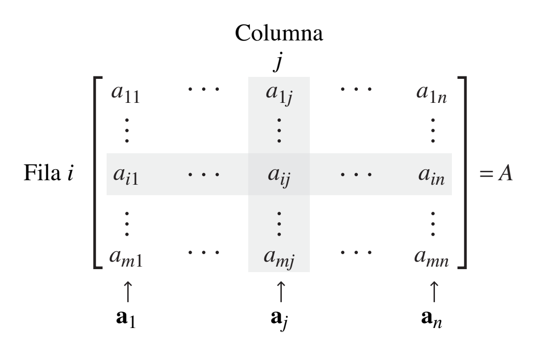


FIGURA 1 Notación matricial.

Las **entradas diagonales** en una matriz $m \times n$ $A = [a_{ij}]$ son $a_{11}, a_{22}, a_{33}, \ldots$, y forman la **diagonal principal** de A. Una **matriz diagonal** es una matriz cuadrada cuyas entradas no diagonales son cero. Un ejemplo es la matriz identidad $n \times n$, I_n. Una matriz de $m \times n$ cuyas entradas son todas cero es una **matriz cero** y se escribe como 0. El tamaño de 0, por lo general, resulta evidente a partir del contexto.

Sumas y múltiplos escalares

La aritmética para vectores que se describió anteriormente admite una extensión natural hacia las matrices. Se dice que dos matrices son **iguales** si tienen el mismo tamaño (es decir, el mismo número de filas y de columnas) y sus columnas correspondientes son iguales, lo cual equivale a decir que sus entradas correspondientes son iguales. Si A y B son matrices $m \times n$, entonces la **suma** $A + B$ es la matriz $m \times n$ cuyas columnas son las sumas de las columnas correspondientes de A y B. Como la suma vectorial de las columnas se realiza por entradas, cada entrada en $A + B$ es la suma de las entradas correspondientes de A y B. La suma $A + B$ está definida sólo cuando A y B son del mismo tamaño.

EJEMPLO 1 Sean

$$A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}, \qquad C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$$

Entonces

$$A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$$

pero $A + C$ no está definida porque A y C tienen diferentes tamaños.

Si r es un escalar y A es una matriz, entonces el **múltiplo escalar** rA es la matriz cuyas columnas son r veces las columnas correspondientes de A. Al igual que con los vectores, se define $-A$ como $(-1)A$ y se escribe $A - B$ en lugar de $A + (-1)B$.

EJEMPLO 2 Si A y B son las matrices del ejemplo 1, entonces

$$2B = 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}$$

$$A - 2B = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix}$$

En el ejemplo 2 no fue necesario calcular $A - 2B$ como $A + (-1)2B$ porque las reglas usuales del álgebra pueden aplicarse a las sumas y múltiplos escalares de matrices, como se verá en el teorema siguiente.

TEOREMA 1

Sean A, B y C matrices del mismo tamaño, y sean r y s escalares.

a. $A + B = B + A$
b. $(A + B) + C = A + (B + C)$
c. $A + 0 = A$
d. $r(A + B) = rA + rB$
e. $(r + s)A = rA + sA$
f. $r(sA) = (rs)A$

Cada igualdad del teorema 1 se verifica mostrando que la matriz del miembro izquierdo tiene el mismo tamaño que la del miembro derecho y que las columnas correspondientes son iguales. El tamaño no es problema porque A, B y C son de igual tamaño. La igualdad de columnas es consecuencia inmediata de las propiedades análogas de los vectores. Por ejemplo, si las columnas j-ésimas de A, B y C son $\mathbf{a}_j$, $\mathbf{b}_j$ y $\mathbf{c}_j$, respectivamente, entonces las columnas j-ésimas de $(A + \mathrm{B}) + C$ y de $A + (B + C)$ son

$$(\mathbf{a}_j + \mathbf{b}_j) + \mathbf{c}_j \qquad \text{y} \qquad \mathbf{a}_j + (\mathbf{b}_j + \mathbf{c}_j)$$

respectivamente. Como estas dos sumas vectoriales son iguales para cada j, la propiedad (b) queda verificada.

Debido a la propiedad asociativa de la suma, es posible escribir simplemente $A + B + C$ para la suma, la cual se puede calcular como $(A + B) + C$ o como $A + (B + C)$. Lo mismo es aplicable para sumas de cuatro o más matrices.

Multiplicación de matrices

Cuando una matriz B multiplica a un vector $\mathbf{x}$, transforma a $\mathbf{x}$ en el vector $B\mathbf{x}$. Si después este vector se multiplica por una matriz A, el vector resultante es $A(B\mathbf{x})$. Vea la figura 2.

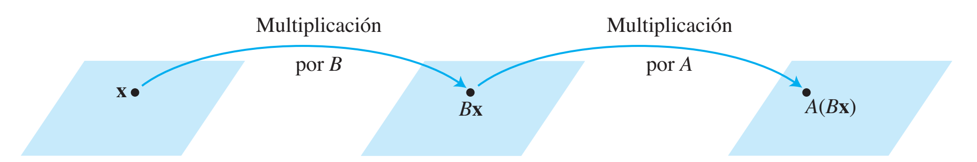


FIGURA 2 Multiplicación por B y luego por A.

Entonces $A(B\mathbf{x})$ se produce a partir de $\mathbf{x}$ gracias a una *composición* de funciones —las transformaciones lineales estudiadas en la sección 1.8. La meta aquí es representar dicha función compuesta como la multiplicación por una matriz única, denotada mediante AB, de manera que

$$A(B\mathbf{x}) = (AB)\mathbf{x} \tag{1}$$

Vea la figura 3.

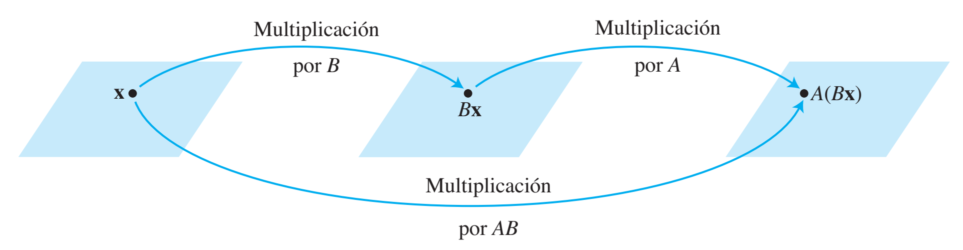


FIGURA 3 Multiplicación por AB.

Si A es de $m \times n$, B es de $n \times p$, y $\mathbf{x}$ está en $\mathbb{R}^p$, denote las columnas de B mediante $\mathbf{b}_1, \ldots, \mathbf{b}_p$, y las entradas de $\mathbf{x}$ mediante $x_1, \ldots, x_p$. Entonces

$$B\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p$$

Por la propiedad de linealidad de la multiplicación por A.

$$\begin{aligned} A(B\mathbf{x}) &= A(x_1\mathbf{b}_1) + \cdots + A(x_p\mathbf{b}_p) \\ &= x_1A\mathbf{b}_1 + \cdots + x_pA\mathbf{b}_p \end{aligned}$$

El vector $A(B\mathbf{x})$ es una combinación lineal de los vectores $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$, usando las entradas de $\mathbf{x}$ como pesos. Al reescribir estos vectores como las columnas de una matriz, se tiene

$$A(B\mathbf{x}) = [\, A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p \,]\,\mathbf{x}$$

Entonces la multiplicación por $[A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]$ transforma a $\mathbf{x}$ en $A(B\mathbf{x})$. ¡Con lo cuál se llega a la matriz buscada!

DEFINICIÓN

Si A es una matriz $m \times n$, y si B es una matriz $n \times p$ con columnas $\mathbf{b}_1, \ldots, \mathbf{b}_p$, entonces el producto AB es la matriz $m \times p$ cuyas columnas son $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$. Esto es,

$$AB = A\,[\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p\,] = [\,A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p\,]$$

Esta definición convierte en verdadera a (1) para toda $\mathbf{x}$ en $\mathbb{R}^p$. La ecuación (1) prueba que la función compuesta de la figura 3 es una transformación lineal y que su matriz estándar es AB. La *multiplicación de matrices corresponde a la composición de transformaciones lineales.*

EJEMPLO 3 Calcule AB, donde $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

Solución Escriba $B = [\mathbf{b}_1 \ \ \mathbf{b}_2 \ \ \mathbf{b}_3]$, y calcule:

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad A\mathbf{b}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad\qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad\qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Entonces

$$AB = A\,[\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3\,] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$$\uparrow \qquad \uparrow \qquad \uparrow$$
$$A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad A\mathbf{b}_3$$

Observe que, como la primer columna de AB es $A\mathbf{b}_1$, esta columna es una combinación lineal de las columnas de A usando como pesos las entradas de $\mathbf{b}_1$. Un enunciado similar es verdadero para cada columna de AB.

Cada columna de AB es una combinación lineal de las columnas de A usando pesos de la columna correspondiente de B.

Desde luego, el número de columnas de A debe corresponder al número de filas que haya en B para que una combinación lineal como $A\mathbf{b}_1$ esté definida. También, la definición de AB muestra que *AB tiene el mismo número de filas que A y el mismo número de columnas que B.*

EJEMPLO 4 Si A es una matriz de 3×5 y B una matriz de 5×2, ¿cuáles son los tamaños de AB y de BA, si tales productos están definidos?

Solución Como A tiene 5 columnas y B tiene 5 filas, el producto AB está definido y es una matriz de 3×2:

$$\underset{3\times 5}{\overset{A}{\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}}} \underset{5\times 2}{\overset{B}{\begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix}}} = \underset{3\times 2}{\overset{AB}{\begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix}}}$$

Corresponden

Tamaño de AB

El producto BA *no* está definido, porque las dos columnas de B no corresponden con las tres filas de A.

La definición de AB es importante para el trabajo teórico y las aplicaciones, pero la siguiente regla proporciona un método más eficiente para calcular las entradas individuales de AB cuando se resuelven a mano problemas sencillos.

REGLA FILA-COLUMNA PARA CALCULAR AB

Si el producto AB está definido, entonces la entrada en la fila i y la columna j de AB es la suma de los productos de entradas correspondientes de la fila i de A y la columna j de B. Si $(AB)_{ij}$ denota la entrada (i, j) en AB, y si A es una matriz $m \times n$, entonces

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Para verificar esta regla, sea $B = [\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_p]$. La columna j de AB es $A\mathbf{b}_j$, y puede calcularse $A\mathbf{b}_j$ por medio de la regla fila-vector para calcular $A\mathbf{x}$ a partir de la sección 1.4. La i-ésima entrada de $A\mathbf{b}_j$ es la suma de los productos de entradas correspondientes de la fila i de A y del vector $\mathbf{b}_j$, que es precisamente el cálculo descrito en la regla para calcular la entrada (i, j) de AB.

EJEMPLO 5 Use la regla fila-columna para calcular dos de las entradas de AB para las matrices del ejemplo 3. Una inspección de los números involucrados aclarará cómo los dos métodos para calcular AB producen la misma matriz.

Solución Para encontrar la entrada de la fila 1 y la columna 3 de AB, considere la fila 1 de A y la columna 3 de B. Multiplique las entradas correspondientes y sume los resultados, como se muestra a continuación:

$$AB = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6) + 3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

Para la entrada en la fila 2 y la columna 2 de AB, use la fila 2 de A y la columna 2 de B:

$$\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3) + -5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$

EJEMPLO 6 Encuentre las entradas de la segunda fila de AB, donde

$$A = \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}, \qquad B = \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

Solución Por aplicación de la regla fila-columna, las entradas de la segunda fila de AB provienen de la fila 2 de A (y las columnas de B):

$$\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}\begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} \square & \square \\ -4 + 21 - 12 & 6 + 3 - 8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

Observe que, como el ejemplo 6 pedía solamente la segunda fila de AB, se podría haber escrito únicamente la segunda fila de A a la izquierda de B y haber calculado

$$\begin{bmatrix} -1 & 3 & -4 \end{bmatrix}\begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 1 \end{bmatrix}$$

Esta observación acerca de las filas de AB es cierta en general, y es consecuencia de la regla fila-columna. Si $\text{fila}_i(A)$ denota la i-ésima fila de una matriz A, entonces

$$\text{fila}_i(AB) = \text{fila}_i(A) \cdot B \qquad (2)$$

Propiedades de la multiplicación de matrices

El teorema siguiente enumera las propiedades estándar de la multiplicación de matrices. Recuerde que I_m representa la matriz identidad $m \times m$, y que $I_m\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^m$.

TEOREMA 2 Sea A una matriz $m \times n$, y sean B y C matrices con tamaños para los cuales las sumas y los productos indicados están definidos.

a. $A(BC) = (AB)C$ (ley asociativa de la multiplicación)

b. $A(B + C) = AB + AC$ (ley distributiva izquierda)

c. $(B + C)A = BA + CA$ (ley distributiva derecha)

d. $r(AB) = (rA)B = A(rB)$ para cualquier escalar r

e. $I_m A = A = AI_n$ (identidad de la multiplicación de matrices)

DEMOSTRACIÓN Las propiedades de la (b) a la (e) se consideran en los ejercicios. La propiedad (a) es consecuencia de que la multiplicación de matrices corresponde a la composición de transformaciones lineales (las cuales son funciones) y es sabido (o fácil de verificar) que la composición de funciones es asociativa. A continuación se presenta otra demostración de (a) que se basa en la “definición de columna” del producto de dos matrices. Sea

$$C = [\,\mathbf{c}_1 \quad \cdots \quad \mathbf{c}_p\,]$$

Por la definición de multiplicación de matrices,

$$BC = [\,B\mathbf{c}_1 \quad \cdots \quad B\mathbf{c}_p\,]$$
$$A(BC) = [\,A(B\mathbf{c}_1) \quad \cdots \quad A(B\mathbf{c}_p)\,]$$

Recuerde de la ecuación (1) que la definición de AB hace que $A(B\mathbf{x}) = (AB)\mathbf{x}$ para toda $\mathbf{x}$, de esta manera

$$A(BC) = [\,(AB)\mathbf{c}_1 \quad \cdots \quad (AB)\mathbf{c}_p\,] = (AB)C \qquad \blacksquare$$

Las leyes asociativa y distributiva de los teoremas 1 y 2 expresan, en esencia, que es posible agregar o quitar parejas de paréntesis en expresiones matriciales de la misma manera que en el álgebra de números reales. En particular, puede escribirse el producto como ABC y calcularlo ya sea como $A(BC)$ o $(AB)C$.[1] De manera similar, se puede calcular un producto de cuatro matrices $ABCD$ como $A(BCD)$ o $(ABC)D$ o $A(BC)D$, y así sucesivamente. No importa cómo se agrupen las matrices al realizar el cálculo de un producto, siempre y cuando se conserve el orden de izquierda a derecha de las matrices.

El orden de izquierda a derecha en productos resulta crítico porque, en general, AB y BA no son iguales. Esto no debe sorprender, porque las columnas de AB son combinaciones lineales de las columnas de A, mientras que las columnas de BA se construyen a partir de las columnas de B. La posición de los factores en el producto AB se enfatiza al decir que A está *multiplicada a la derecha* por B o que B está *multiplicada a la izquierda* por A. Si $AB = BA$, se dice que A y B **conmutan** una con la otra.

[1]Cuando B es cuadrada y C tiene menos columnas que las filas que tiene A, resulta más eficiente calcular $A(BC)$ en lugar de $(AB)C$.

EJEMPLO 7 Sea $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$. Muestre que estas matrices no conmutan. Esto es, verifique que $AB \neq BA$.

Solución

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$

Para enfatizar, se incluye la observación acerca de la conmutatividad con la siguiente lista de diferencias importantes entre el álgebra de matrices y el álgebra de números reales. Si desea ver ejemplos de estas situaciones, consulte los ejercicios del 9 al 12.

Advertencias:

1. En general, $AB \neq BA$.
2. Las leyes de la cancelación *no* se aplican en la multiplicación de matrices. Esto es, si $AB = AC$, en general *no* es cierto que $B = C$. (Vea el ejercicio 10.)
3. Si un producto AB es la matriz cero, en general *no se puede* concluir que $A = 0$ o $B = 0$. (Vea el ejercicio 12.)

Potencias de una matriz

Si A es una matriz $n \times n$ y k es un entero positivo, entonces A^k denota el producto de k copias de A:

$$A^k = \underbrace{A \cdots A}_{k}$$

Si A es distinta de cero y si $\mathbf{x}$ está en $\mathbb{R}^n$, entonces $A^k\mathbf{x}$ es el resultado de multiplicar $\mathbf{x}$ repetidamente a la izquierda por A, k veces. Si $k = 0$, entonces $A^0\mathbf{x}$ debe ser la misma $\mathbf{x}$. Por lo tanto, A^0 se interpreta como la matriz identidad. Las potencias de matrices son útiles tanto en la teoría como en las aplicaciones (secciones 2.6, 4.9, y posteriormente en el texto).

La transpuesta de una matriz

Dada una matriz A de $m \times n$, la **transpuesta** de A es la matriz $n \times m$, denotada mediante A^T, cuyas columnas se forman a partir de las filas correspondientes de A.

EJEMPLO 8 Sean

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

Entonces

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \qquad B^T = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \qquad C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$

TEOREMA 3

Sean A y B matrices cuyos tamaños son apropiados para las sumas y los productos siguientes.

a. $(A^T)^T = A$

b. $(A + B)^T = A^T + B^T$

c. Para cualquier escalar r, $(rA)^T = rA^T$

d. $(AB)^T = B^T A^T$

Las demostraciones de (a) a (c) son directas y se omiten. Para (d), vea el ejercicio 33. Por lo general, $(AB)^T$ no es igual a $A^T B^T$, aún cuando A y B tengan tamaños tales que el producto $A^T B^T$ esté definido.

La generalización del teorema 3(d) a productos de más de dos factores puede establecerse en palabras de la manera siguiente:

La transpuesta de un producto de matrices es igual al producto de sus transpuestas en el orden *inverso*.

Los ejercicios contienen ejemplos numéricos que ilustran las propiedades de las transpuestas.

Notas numéricas

1. La manera más rápida de obtener AB en una computadora depende de la forma en que la computadora guarde las matrices en su memoria. Los algoritmos estándar de mayor eficiencia, tales como los de LAPACK, calculan AB por columnas, como en la definición del producto presentada en este texto. (Una versión de LAPACK escrita en C++ calcula AB por filas.)
2. La definición de AB se presta al procesamiento paralelo en una computadora. Las columnas de B se asignan individualmente o en grupos a diferentes procesadores, los cuales de manera independiente y, por lo tanto, simultánea calculan las columnas correspondientes de AB.

PROBLEMAS DE PRÁCTICA

1. Dado que los vectores en $\mathbb{R}^n$ pueden verse como matrices $n \times 1$, las propiedades de las transpuestas del teorema 3 también se aplican a vectores. Sean

$$A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

Calcule $(A\mathbf{x})^T$, $\mathbf{x}^T A^T$, $\mathbf{x}\mathbf{x}^T$, y $\mathbf{x}^T\mathbf{x}$. ¿Está definida $A^T\mathbf{x}^T$?

2. Sean A una matriz de 4×4 y $\mathbf{x}$ un vector en $\mathbb{R}^4$. ¿Cuál es la forma más rápida de calcular $A^2\mathbf{x}$? Cuente las multiplicaciones.

2.1 EJERCICIOS

En los ejercicios 1 y 2, calcule cada suma o producto si la matriz está definida. Si alguna expresión no está definida, explique por qué. Sean

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & -5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & -5 & 1 \\ 1 & -4 & -3 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 5 \\ -1 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

1. $-2A$, $B - 2A$, AC, CD

2. $A + 2B$, $3C - E$, CB, EB

En el resto de esta serie de ejercicios y en las series que siguen, debe suponerse que cada expresión de matrices está definida. Esto es, los tamaños de las matrices (y los vectores) involucrados "se corresponden" de manera apropiada.

3. Sea $A = \begin{bmatrix} 4 & -1 \\ 5 & -2 \end{bmatrix}$. Calcule $3I_2 - A$ y $(3I_2)A$.

4. Calcule $A - 5I_3$ y $(5I_3)A$, cuando

$$A = \begin{bmatrix} 9 & -1 & 3 \\ -8 & 7 & -6 \\ -4 & 1 & 8 \end{bmatrix}.$$

En los ejercicios 5 y 6, calcule el producto AB en dos formas: (a) mediante la definición, donde $A\mathbf{b}_1$ y $A\mathbf{b}_2$ se calculan por separado, y (b) mediante la regla fila-columna para calcular AB.

5. $A = \begin{bmatrix} -1 & 2 \\ 5 & 4 \\ 2 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & -2 \\ -2 & 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 4 & -2 \\ -3 & 0 \\ 3 & 5 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$

7. Si una matriz A es de 5×3 y el producto AB es de 5×7, ¿cuál es el tamaño de B?

8. ¿Cuántas filas tiene B si BC es una matriz de 3×4?

9. Sean $A = \begin{bmatrix} 2 & 5 \\ -3 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & -5 \\ 3 & k \end{bmatrix}$. ¿Qué valor(es) de k, si hay, hacen que $AB = BA$?

10. Sean $A = \begin{bmatrix} 2 & -3 \\ -4 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 8 & 4 \\ 5 & 5 \end{bmatrix}$, y $C = \begin{bmatrix} 5 & -2 \\ 3 & 1 \end{bmatrix}$.

Verifique que $AB = AC$ y que sin embargo $B \neq C$.

11. Sean $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 5 \end{bmatrix}$ y $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$. Calcule AD y DA. Explique cómo cambian las filas o columnas de A cuando se multiplica por D a la derecha o a la izquierda. Encuentre una matriz B de 3×3, que no sea la matriz identidad o la matriz cero, tal que $AB = BA$.

12. Sea $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$. Construya una matriz B de 2×2 tal que AB sea igual a la matriz cero. Las columnas de B no deben ser iguales entre sí y deben ser distintas de cero.

13. Sean $\mathbf{r}_1, \ldots, \mathbf{r}_p$ vectores en $\mathbb{R}^n$, y sea Q una matriz $m \times n$. Escriba la matriz $[Q\mathbf{r}_1 \ \cdots \ Q\mathbf{r}_p]$ como un *producto* de dos matrices (ninguna de ellas igual a la matriz identidad).

14. Sea U la matriz de 3×2 de costos descrita en el ejemplo 6 de la sección 1.8. La primera columna de U enlista los costos por dólar de producción para elaborar el producto B, y la segunda columna enlista los costos por dólar de producción para el artículo C. (Los costos tienen las categorías de materiales, mano de obra, y gastos generales.) Sea $\mathbf{q}_1$ un vector en $\mathbb{R}^2$ que enlista la producción (medida en dólares) de los bienes B y C fabricados durante el primer trimestre del año, y sean $\mathbf{q}_2$, $\mathbf{q}_3$ y $\mathbf{q}_4$ los vectores análogos que muestran las cantidades de producto B y C fabricadas en el segundo, tercero y cuarto trimestre, respectivamente. Proporcione una descripción económica de los datos en la matriz UQ, donde $Q = [\mathbf{q}_1 \ \ \mathbf{q}_2 \ \ \mathbf{q}_3 \ \ \mathbf{q}_4]$.

Los ejercicios 15 y 16 tratan de matrices arbitrarias A, B y C para las cuales las sumas y productos indicados están definidos. Señale cada afirmación como verdadera o falsa. Justifique sus respuestas.

15. a. Si A y B son de 2×2 con columnas $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{b}_1$, $\mathbf{b}_2$, respectivamente, entonces $AB = [\mathbf{a}_1\mathbf{b}_1 \quad \mathbf{a}_2\mathbf{b}_2]$.

b. Toda columna de AB es una combinación lineal de las columnas de B usando pesos de la columna correspondiente de A.

c. $AB + AC = A(B + C)$

d. $A^T + B^T = (A + B)^T$

e. La transpuesta de un producto de matrices es igual al producto de sus transpuestas en el mismo orden.

16. a. Si A y B son de 3×3 y $B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]$, entonces $AB = [A\mathbf{b}_1 + A\mathbf{b}_2 + A\mathbf{b}_3]$.

b. La segunda fila de AB es la segunda fila de A multiplicada a la derecha por B.

c. $(AB)C = (AC)B$

d. $(AB)^T = A^TB^T$

e. La transpuesta de una suma de matrices es igual a la suma de sus transpuestas.

17. Si $A = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$ y $AB = \begin{bmatrix} -1 & 2 & -1 \\ 6 & -9 & 3 \end{bmatrix}$, determine la primera y la segunda columnas de B.

18. Suponga que las dos primeras columnas de B, $\mathbf{b}_1$ y $\mathbf{b}_2$ son iguales. ¿Qué puede decirse acerca de las columnas de AB (si AB está definida)? ¿Por qué?

19. Suponga que la tercera columna de B es la suma de las primeras dos columnas. ¿Qué puede decirse acerca de la tercera columna de AB? ¿Por qué?

20. Suponga que la segunda columna de B es toda cero. ¿Qué puede decirse acerca de la segunda columna de AB?

21. Suponga que la última columna de AB es completamente cero, pero B por sí sola no tiene ninguna columna de ceros. ¿Qué puede decirse acerca de las columnas de A?

22. Muestre que si las columnas de B son linealmente dependientes, también lo son las columnas de AB.

23. Suponga que $CA = I_n$ (la matriz identidad $n \times n$). Muestre que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial. Explique por qué A no puede tener más columnas que filas.

24. Suponga que $AD = I_m$ (la matriz identidad $m \times m$). Muestre que para toda $\mathbf{b}$ en $\mathbb{R}^m$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución. [*Sugerencia:* Piense en la ecuación $AD\mathbf{b} = \mathbf{b}$.] Explique por qué A no puede tener más filas que columnas.

25. Suponga que A es una matriz $m \times n$ y que existen matrices $n \times m$, C y D, tales que $CA = I_n$ y $AD = I_m$. Pruebe que $m = n$ y $C = D$. [*Sugerencia:* Piense en el producto CAD.]

26. Suponga que A es una matriz de $3 \times n$ cuyas columnas generan $\mathbb{R}^3$. Explique cómo construir una matriz D de $n \times 3$ tal que $AD = I_3$.

En los ejercicios 27 y 28, vea los vectores en $\mathbb{R}^n$ como matrices $n \times 1$. Para $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, el producto de matrices $\mathbf{u}^T\mathbf{v}$ es una matriz 1×1, llamada **producto escalar**, o **producto interno**, de $\mathbf{u}$ y $\mathbf{v}$. Por lo general, se escribe como un único número real sin corchetes. El producto de matrices $\mathbf{u}\mathbf{v}^T$ es una matriz $n \times n$, llamada **producto exterior** de $\mathbf{u}$ y $\mathbf{v}$. Los productos $\mathbf{u}^T\mathbf{v}$ y $\mathbf{u}\mathbf{v}^T$ aparecerán más adelante en el texto.

27. Sean $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \\ -4 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Calcule $\mathbf{u}^T\mathbf{v}$, $\mathbf{v}^T\mathbf{u}$, $\mathbf{u}\mathbf{v}^T$ y $\mathbf{v}\mathbf{u}^T$.

28. Si $\mathbf{u}$ y $\mathbf{v}$ están en $\mathbb{R}^n$, ¿qué relación hay entre $\mathbf{u}^T\mathbf{v}$ y $\mathbf{v}^T\mathbf{u}$? ¿Y entre $\mathbf{u}\mathbf{v}^T$ y $\mathbf{v}\mathbf{u}^T$?

29. Compruebe el teorema 2(b) y 2(c). Use la regla fila-columna. La entrada (i, j) de $A(B + C)$ se puede escribir como

$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj}) \text{ o bien } \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

30. Compruebe el teorema 2(d). [*Sugerencia:* La entrada (i, j) en $(rA)B$ es $(ra_{i1})b_{1j} + \cdots + (ra_{in})b_{nj}$.]

31. Muestre que $I_mA = A$ cuando A es una matriz $m \times n$. Se puede suponer que $I_m\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^m$.

32. Muestre que $AI_n = A$ cuando A es una matriz $m \times n$. [*Sugerencia:* Use la definición (de columna) de AI_n.]

33. Compruebe el teorema 3(d). [*Sugerencia:* Considere la j-ésima fila de $(AB)^T$.]

34. Proporcione una fórmula para $(AB\mathbf{x})^T$, donde $\mathbf{x}$ es un vector y A y B son matrices con los tamaños apropiados.

35. **[M]** Lea la documentación de su programa de matrices y escriba los comandos que producirían las siguientes matrices (sin introducir cada entrada de la matriz).

a. Una matriz de ceros de 5×6.

b. Una matriz de unos de 3×5.

c. La matriz identidad de 6×6.

d. Una matriz diagonal de 5×5, con entradas diagonales 3, 5, 7, 2, 4.

Una forma útil de probar ideas nuevas o de formular conjeturas en álgebra de matrices es realizar cálculos con matrices seleccionadas en forma aleatoria. La comprobación de una propiedad para unas cuantas matrices no demuestra que la propiedad sea válida en general, pero permite que la propiedad sea más creíble. Además, si una propiedad es falsa, esto puede descubrirse al realizar unos cuantos cálculos.

36. **[M]** Escriba el comando o los comandos necesarios para crear una matriz de 6×4 con entradas al azar. ¿Dentro de qué rango de números están las entradas? Diga cómo crear

una matriz aleatoria de 3×3 con entradas enteras entre -9 y 9. [*Sugerencia:* Si x es un número aleatorio tal que $0 < x < 1$, entonces $-9.5 < 19(x - 0.5) < 9.5$.]

37. **[M]** Construya una matriz aleatoria A de 4×4 y compruebe si $(A + I)(A - I) = A^2 - I$. La mejor manera de hacer esto es calcular $(A + I)(A - I) - (A^2 - I)$, y verificar que esta diferencia sea la matriz cero. Hágalo para tres matrices al azar. Luego realice la prueba para $(A + B)(A - B) = A^2 - B^2$, procediendo en la misma forma con tres pares de matrices aleatorias de 4×4. Informe las conclusiones obtenidas.

38. **[M]** Use al menos tres pares de matrices aleatorias A y B de 4×4 para probar las igualdades $(A + B)^T = A^T + B^T$ y $(AB)^T = A^TB^T$. (Vea el ejercicio 37.) Informe las conclusiones obtenidas. [*Nota:* La mayoría de los programas de matrices usan A' para representar A^T.]

39. **[M]** Sea

$$S = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Calcule S^k para $k = 2, \ldots, 6$.

40. **[M]** Describa con palabras qué pasa al calcular A^5, A^{10}, A^{20} y A^{30} para

$$A = \begin{bmatrix} 1/6 & 1/2 & 1/3 \\ 1/2 & 1/4 & 1/4 \\ 1/3 & 1/4 & 5/12 \end{bmatrix}$$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. $A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$. Así que $(A\mathbf{x})^T = [-4 \quad 2]$. También, $\mathbf{x}^TA^T = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -4 & 2 \end{bmatrix}$. Las cantidades $(A\mathbf{x})^T$ y $\mathbf{x}^TA^T$ son iguales, como cabe esperar por el teorema 3(d). Enseguida,

$$\mathbf{x}\mathbf{x}^T = \begin{bmatrix} 5 \\ 3 \end{bmatrix}\begin{bmatrix} 5 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 15 \\ 15 & 9 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{x} = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = [\,25 + 9\,] = 34$$

Una matriz de 1×1 como $\mathbf{x}^T\mathbf{x}$ generalmente se escribe sin corchetes. Por último, $A^T\mathbf{x}^T$ no está definida, porque $\mathbf{x}^T$ no tiene dos filas que correspondan a las dos columnas de A^T.

2. La manera más rápida de calcular $A^2\mathbf{x}$ es determinando $A(A\mathbf{x})$. El producto $A\mathbf{x}$ requiere 16 multiplicaciones, 4 por cada entrada, y $A(A\mathbf{x})$ requiere 16 más. Por contraste, el producto A^2 requiere 64 multiplicaciones, 4 por cada una de las 16 entradas en A^2. Después de eso, $A^2\mathbf{x}$ requiere 16 multiplicaciones más, para un total de 80.

2.2 LA INVERSA DE UNA MATRIZ

El álgebra de matrices proporciona herramientas para manipular ecuaciones matriciales y crear diversas fórmulas útiles en formas similares a la ejecución ordinaria del álgebra con números reales. En esta sección se investiga el análogo matricial del recíproco, o inverso multiplicativo, de un número diferente de cero.

Recuerde que el inverso multiplicativo de un número como 5 es $1/5$ o 5^{-1}. Este inverso satisface la ecuación

$$5^{-1} \cdot 5 = 1 \quad \text{y} \quad 5 \cdot 5^{-1} = 1$$

La generalización matricial requiere *ambas* ecuaciones y evita la notación con diagonales (para indicar una división) debido a que la multiplicación de matrices no es conmutativa. Más aún, una generalización completa sólo es posible si las matrices involucradas son cuadradas.[1]

Se dice que una matriz A de $n \times n$ es **invertible** si existe otra matriz C de $n \times n$ tal que

$$CA = I \quad \text{y} \quad AC = I$$

donde $I = I_n$, la matriz identidad $n \times n$. En este caso, C es un **inverso** de A. De hecho, C está determinado únicamente por A, porque si B fuera otro inverso de A, entonces $B = BI = B(AC) = (BA)C = IC = C$. Este inverso único se denota mediante A^{-1}, de manera que,

$$A^{-1}A = I \quad \text{y} \quad AA^{-1} = I$$

Una matriz que *no* es invertible algunas veces se denomina **matriz singular**, y una matriz invertible se denomina **matriz no singular**.

EJEMPLO 1 Si $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ y $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, entonces

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y}$$

$$CA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Así que $C = A^{-1}$.

A continuación se presenta una fórmula sencilla para el inverso de una matriz de 2×2, junto con una prueba para saber si existe el inverso.

TEOREMA 4

Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Si $ad - bc \neq 0$, entonces A es invertible y

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Si $ad - bc = 0$, entonces A no es invertible.

La demostración sencilla del teorema 4 se describe en términos generales en los ejercicios 25 y 26. La cantidad $ad - bc$ se llama **determinante** de A, y se escribe

$$\det A = ad - bc$$

El teorema 4 establece que una matriz A de 2×2 es invertible si, y sólo si, $\det A \neq 0$.

[1]Podría decirse que una matriz A de $m \times n$ es invertible si existen matrices $n \times m$, C y D, tales que $CA = I_n$ y $AD = I_m$. Sin embargo, estas ecuaciones implican que A es cuadrada y $C = D$. Por lo tanto, A es invertible como se definió con anterioridad. Vea los ejercicios 23, 24 y 25 en la sección 2.1.

EJEMPLO 2 Encuentre el inverso de $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$.

Solución Como $\det A = 3(6) - 4(5) = -2 \neq 0$, A es invertible, y

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 6/(-2) & -4/(-2) \\ -5/(-2) & 3/(-2) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}$$ ■

Las matrices invertibles son indispensables en el álgebra lineal —principalmente para cálculos algebraicos y deducciones de fórmulas, como en el teorema siguiente. En ocasiones una matriz inversa permite entender mejor un modelo matemático de alguna situación de la vida real, como en el ejemplo 3 que se presenta más adelante.

TEOREMA 5

Si A es una matriz invertible $n \times n$ entonces, para cada $\mathbf{b}$ en $\mathbb{R}^n$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$.

DEMOSTRACIÓN Tome cualquier $\mathbf{b}$ en $\mathbb{R}^n$. Existe una solución porque cuando se sustituye $A^{-1}\mathbf{b}$ por $\mathbf{x}$, se tiene $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$. Así que $A^{-1}\mathbf{b}$ es una solución. Para probar que la solución es única, se muestra que si $\mathbf{u}$ es cualquier solución, entonces $\mathbf{u}$ debe ser, de hecho, $A^{-1}\mathbf{b}$. En efecto, si $A\mathbf{u} = \mathbf{b}$, pueden multiplicarse ambos miembros por A^{-1} y obtener

$$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}, \qquad I\mathbf{u} = A^{-1}\mathbf{b}, \qquad \text{y} \qquad \mathbf{u} = A^{-1}\mathbf{b}$$ ■

EJEMPLO 3 Una viga elástica horizontal tiene soportes en cada extremo y está sometida a fuerzas en los puntos 1, 2, 3, como indica la figura 1. Sea $\mathbf{f}$ en $\mathbb{R}^3$ tal que enliste las fuerzas en estos puntos, y sea $\mathbf{y}$ en $\mathbb{R}^3$ tal que incluya las magnitudes de la deflexión (esto es, movimiento) de la viga en los tres puntos. Al aplicar la ley de Hooke de la física, se puede demostrar que

$$\mathbf{y} = D\mathbf{f}$$

donde D es una *matriz de flexibilidad.* Su inversa se denomina *matriz de rigidez.* Describa el significado físico de las columnas de D y D^{-1}.

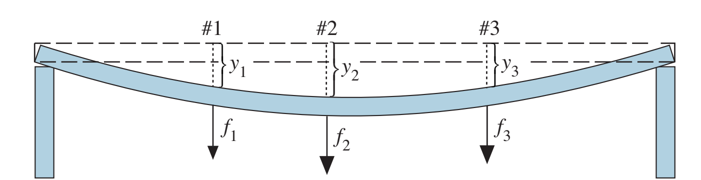


FIGURA 1 Deflexión de una viga elástica.

Solución Escriba $I_3 = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3]$ y observe que

$$D = DI_3 = [D\mathbf{e}_1 \quad D\mathbf{e}_2 \quad D\mathbf{e}_3]$$

Interprete el vector $\mathbf{e}_1 = (1, 0, 0)$ como una fuerza unitaria aplicada hacia abajo en el punto 1 (con fuerza cero en los otros dos puntos). Entonces la primera columna de D,

$D\mathbf{e}_1$, enlista las deflexiones debidas a una fuerza unitaria en el punto 1. Interpretaciones similares son válidas para la segunda y tercera columnas de D.

Para estudiar la matriz de rigidez D^{-1}, observe que la ecuación $\mathbf{f} = D^{-1}\mathbf{y}$ calcula un vector de fuerza $\mathbf{f}$ cuando se da un vector de deflexión $\mathbf{y}$. Escriba

$$D^{-1} = D^{-1}I_3 = [D^{-1}\mathbf{e}_1 \quad D^{-1}\mathbf{e}_2 \quad D^{-1}\mathbf{e}_3]$$

Ahora interprete $\mathbf{e}_1$ como un vector de deflexión. Entonces $D^{-1}\mathbf{e}_1$ enlista las fuerzas que crean la deflexión. Esto es, la primera columna de D^{-1} enlista las fuerzas que deben aplicarse en los tres puntos para producir una deflexión unitaria en el punto 1 y cero deflexión en los otros puntos. De manera similar, las columnas 2 y 3 de D^{-1} enlistan las fuerzas requeridas para producir deflexiones unitarias en los puntos 2 y 3, respectivamente. En cada columna, una o dos de las fuerzas deben ser negativas (apuntar hacia arriba) para producir una deflexión unitaria en el punto deseado y cero deflexión en los otros dos puntos. Si la flexibilidad se mide, por ejemplo, en pulgadas de deflexión por libra de carga, entonces las entradas de la matriz de rigidez están dadas en libras de carga por pulgada de deflexión.

La fórmula del teorema 5 se utiliza muy pocas veces para resolver en forma numérica una ecuación $A\mathbf{x} = \mathbf{b}$ porque la reducción por filas de $[A \quad \mathbf{b}]$ casi siempre es más rápida. (La reducción por filas es también más precisa, generalmente, cuando los cálculos requieren el redondeo de los números.) Una posible excepción es el caso de 2×2; ya que los cálculos mentales para resolver $A\mathbf{x} = \mathbf{b}$ en ocasiones resultan más fáciles usando la fórmula para A^{-1}, como en el siguiente ejemplo.

EJEMPLO 4 Use el inverso de la matriz A del ejemplo 2 para resolver el sistema

$$\begin{aligned} 3x_1 + 4x_2 &= 3 \\ 5x_1 + 6x_2 &= 7 \end{aligned}$$

Solución Este sistema es equivalente a $A\mathbf{x} = \mathbf{b}$, así que

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

El teorema siguiente proporciona tres datos útiles acerca de las matrices invertibles.

TEOREMA 6

a. Si A es una matriz invertible, entonces A^{-1} es invertible y

$$(A^{-1})^{-1} = A$$

b. Si A y B son matrices invertibles de $n \times n$, entonces también lo es AB, y el inverso de AB es el producto de los inversos de A y B en el orden opuesto. Esto es,

$$(AB)^{-1} = B^{-1}A^{-1}$$

c. Si A es una matriz invertible, también lo es A^T, y el inverso de A^T es la transpuesta de A^{-1}. Esto es,

$$(A^T)^{-1} = (A^{-1})^T$$

DEMOSTRACIÓN Para verificar (a), debe encontrarse una matriz C tal que

$$A^{-1}C = I \quad \text{y} \quad CA^{-1} = I$$

Sin embargo, ya se sabe que estas ecuaciones se satisfacen colocando a A en lugar de C. Por lo tanto, A^{-1} es invertible y A es su inverso. Enseguida, para demostrar (b), se calcula:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

Un cálculo similar muestra que $(B^{-1}A^{-1})(AB) = I$. Para (c) es aplicable el teorema 3(d), lea de izquierda a derecha, $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$. De manera similar, $A^T(A^{-1})^T = I^T = I$. Por lo tanto, A^T es invertible, y su inverso es $(A^{-1})^T$. ■

La siguiente generalización del teorema 6(b) se necesitará más adelante.

El producto de matrices invertibles de $n \times n$ es invertible, y el inverso es el producto de sus inversos en el orden opuesto.

Existe una conexión importante entre las matrices invertibles y las operaciones de fila que conduce a un método para calcular inversos. Como se verá, una matriz invertible A es equivalente por filas a una matriz identidad, y se puede encontrar A^{-1} al *observar la reducción por filas de A a I.*

Matrices elementales

Una **matriz elemental** es aquella que se obtiene al realizar una única operación elemental de fila sobre una matriz identidad. El siguiente ejemplo ilustra los tres tipos de matrices elementales.

EJEMPLO 5 Sean

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix},$$

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Calcule E_1A, E_2A y E_3A, y describa cómo se pueden obtener estos productos por medio de operaciones elementales de fila sobre A.

Solución Se tiene

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, \qquad E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix},$$

$$E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}$$

La suma de −4 veces la fila 1 de A a la fila 3 produce E_1A. (Ésta es una operación de reemplazo de fila.) Un intercambio de las filas 1 y 2 de A produce E_2A, y la multiplicación de la fila 3 de A por 5 produce E_3A.

La multiplicación izquierda (esto es, multiplicación por la izquierda) por E_1 en el ejemplo 5 tiene el mismo efecto en cualquier matriz de $3 \times n$. Esta multiplicación suma −4 veces la fila 1 a la fila 3. En particular, como $E_1 \cdot I = E_1$, se observa que la *misma* E_1 se produce por medio de esta misma operación de fila sobre la identidad. Así, el ejemplo 5 ilustra la siguiente propiedad general de las matrices elementales. Vea los ejercicios 27 y 28.

Si se realiza una operación elemental de fila con una matriz A de $m \times n$, la matriz resultante puede escribirse como EA, donde la matriz E de $m \times m$ se crea al realizar la misma operación de fila sobre I_m.

Debido a que las operaciones de fila son reversibles, como se mostró en la sección 1.1, las matrices elementales son invertibles, porque si E se produce aplicando una operación de fila sobre I, entonces existe otra operación de fila del mismo tipo que convierte a E de nuevo en I. Por lo tanto, existe una matriz elemental F tal que $FE = I$. También, como E y F corresponden a operaciones inversas, $EF = I$.

Toda matriz elemental E es invertible. El inverso de E es la matriz elemental del mismo tipo que transforma a E de nuevo en I.

EJEMPLO 6 Encuentre el inverso de $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

Solución Para transformar E_1 en I, sume +4 veces la fila 1 a la fila 3. La matriz elemental que hace esto es

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$$

El teorema siguiente ofrece la mejor manera de “visualizar” una matriz invertible, y conduce de inmediato a un método para encontrar la inversa de una matriz.

TEOREMA 7 Una matriz A de $n \times n$ es invertible si, y sólo si, A es equivalente por filas a I_n, y en este caso, cualquier secuencia de operaciones elementales de fila que reduzca A a I_n también transforma I_n en A^{-1}.

DEMOSTRACIÓN Suponga que A es invertible. Entonces, como la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para toda $\mathbf{b}$ (teorema 5), A tiene una posición pivote en cada fila (teorema 4 de la sección 1.4). Como A es cuadrada, las n posiciones pivote deben estar sobre la diagonal, lo cual implica que la forma escalonada reducida de A es I_n. Esto es, $A \sim I_n$.

De manera inversa, suponga ahora que $A \sim I_n$. Entonces, puesto que cada paso de la reducción por filas de A corresponde a una multiplicación izquierda por una matriz elemental, existen matrices elementales $E_1, \ldots, E_p$ tales que

$$A \sim E_1 A \sim E_2(E_1 A) \sim \cdots \sim E_p(E_{p-1} \cdots E_1 A) = I_n$$

Esto es,

$$E_p \cdots E_1 A = I_n \tag{1}$$

Puesto que el producto $E_p \cdots E_1$ de matrices invertibles es invertible, (1) conduce a

$$\begin{aligned}(E_p \cdots E_1)^{-1}(E_p \cdots E_1)A &= (E_p \cdots E_1)^{-1} I_n \\ A &= (E_p \cdots E_1)^{-1}\end{aligned}$$

Por lo tanto, A es invertible, porque es el inverso de una matriz invertible (teorema 6). También,

$$A^{-1} = [\,(E_p \cdots E_1)^{-1}\,]^{-1} = E_p \cdots E_1$$

Entonces $A^{-1} = E_p \cdots E_1 \cdot I_n$, lo cual establece que al aplicar sucesivamente $E_1, \ldots, E_p$ a I_n se obtiene A^{-1}. Ésta es la misma secuencia de (1) que redujo A a I_n. ■

Un algoritmo para encontrar A^{-1}

Si se colocan lado a lado A e I para formar una matriz aumentada $[A \quad I]$, entonces las operaciones de fila en esta matriz producen operaciones idénticas sobre A e I. Por el teorema 7, o hay operaciones de fila que transforman a A en I_n y a I_n en A^{-1}, o A no es invertible.

> ALGORITMO PARA ENCONTRAR A^{-1}
>
> Reduzca por filas la matriz aumentada $[A \quad I]$. Si A es equivalente por filas a I, entonces $[A \quad I]$ es equivalente por filas a $[I \quad A^{-1}]$. Si no es así, A no tiene inversa.

EJEMPLO 7 Encuentre el inverso de la matriz $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, si existe.

Solución

$$[\,A \quad I\,] = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

Como $A \sim I$, por el teorema 7 se concluye que A es invertible, y

$$A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$$

Es recomendable verificar la respuesta final:

$$AA^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

No es necesario verificar que $A^{-1}A = I$ puesto que A es invertible. ■

Otro enfoque de inversión de matrices

Denote las columnas de I_n mediante $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Entonces la reducción por filas de $[A \quad I]$ a $[I \quad A^{-1}]$ puede verse como la solución simultánea de los n sistemas

$$A\mathbf{x} = \mathbf{e}_1, \quad A\mathbf{x} = \mathbf{e}_2, \quad \ldots, \quad A\mathbf{x} = \mathbf{e}_n \tag{2}$$

donde todas las “columnas aumentadas” de estos sistemas se han colocado contiguas a A para formar $[A \quad \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n] = [A \quad I]$. La ecuación $AA^{-1} = I$, así como la definición de multiplicación de matrices, muestran que las columnas de A^{-1} son precisamente las soluciones de los sistemas de (2). Esta observación es útil porque en algunos problemas aplicados podría ser necesario encontrar solamente una o dos columnas de A^{-1}. En este caso, basta con resolver los sistemas correspondientes de (2).

NOTA NUMÉRICA

En la práctica, rara vez se calcula A^{-1}, a menos que se necesiten las entradas de A^{-1}. Calcular tanto A^{-1} como $A^{-1}\mathbf{b}$ requiere aproximadamente tres veces más operaciones aritméticas que resolver mediante reducción por filas $A\mathbf{x} = \mathbf{b}$, y la reducción por filas puede resultar más precisa.

CD Exploración de las propiedades de los inversos (Exploring Properties of Inverses)

PROBLEMAS DE PRÁCTICA

1. Utilice determinantes para establecer cuáles de las siguientes matrices son invertibles:

a. $\begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix}$ b. $\begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix}$ c. $\begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix}$

2. Si existe, encuentre el inverso de la matriz $A = \begin{bmatrix} 1 & -2 & -1 \\ -1 & 5 & 6 \\ 5 & -4 & 5 \end{bmatrix}$,

2.2 EJERCICIOS

En los ejercicios 1 a 4 encuentre los inversos de las matrices.

1. $\begin{bmatrix} 8 & 6 \\ 5 & 4 \end{bmatrix}$ **2.** $\begin{bmatrix} 3 & 2 \\ 7 & 4 \end{bmatrix}$

3. $\begin{bmatrix} 8 & 5 \\ -7 & -5 \end{bmatrix}$ **4.** $\begin{bmatrix} 3 & -4 \\ 7 & -8 \end{bmatrix}$

5. Use el inverso encontrado en el ejercicio 1 para resolver el sistema

$$\begin{aligned} 8x_1 + 6x_2 &= 2 \\ 5x_1 + 4x_2 &= -1 \end{aligned}$$

6. Use el inverso encontrado en el ejercicio 3 para resolver el sistema

$$\begin{aligned} 8x_1 + 5x_2 &= -9 \\ -7x_1 - 5x_2 &= 11 \end{aligned}$$

7. Sean $A = \begin{bmatrix} 1 & 2 \\ 5 & 12 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$, y $\mathbf{b}_4 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

a. Encuentre A^{-1} y utilícelo para resolver las cuatro ecuaciones

$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad A\mathbf{x} = \mathbf{b}_3, \quad A\mathbf{x} = \mathbf{b}_4$

b. Las cuatro ecuaciones del inciso (a) pueden resolverse con el *mismo* conjunto de operaciones de fila, puesto que la matriz de coeficientes es la misma en cada caso. Resuelva las cuatro ecuaciones del inciso (a) reduciendo por filas la matriz aumentada $[A \quad \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad \mathbf{b}_4]$.

8. Utilice álgebra de matrices para mostrar que si A es invertible y D satisface $AD = I$, entonces $D = A^{-1}$.

En los ejercicios 9 y 10, señale cada afirmación como verdadera o falsa. Justifique sus respuestas.

9. a. Para que una matriz B sea inverso de A, ambas ecuaciones $AB = I$ y $BA = I$ deben ser ciertas.

b. Si A y B son de $n \times n$ e invertibles, entonces $A^{-1}B^{-1}$ es el inverso de AB.

c. Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ y $ab - cd \neq 0$, entonces A es invertible.

d. Si A es una matriz invertible $n \times n$, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para *toda* $\mathbf{b}$ en $\mathbb{R}^n$.

e. Toda matriz elemental es invertible.

10. a. Un producto de matrices invertibles de $n \times n$ es invertible, y el inverso del producto es el producto de sus inversos en el mismo orden.

b. Si A es invertible, entonces el inverso de A^{-1} es la propia A.

c. Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, y $ad = bc$, entonces A no es invertible.

d. Si A se puede reducir por filas a la matriz identidad, entonces A debe ser invertible.

e. Si A es invertible, entonces las operaciones elementales de fila que reducen A a la identidad I_n también reducen A^{-1} a I_n.

11. Sea A una matriz invertible de $n \times n$ y sea B una matriz $n \times p$. Muestre que la ecuación $AX = B$ tiene una única solución $A^{-1}B$.

12. Sea A una matriz invertible $n \times n$, y sea B una matriz $n \times p$. Explique por qué $A^{-1}B$ puede calcularse mediante reducción por filas:

Si $[A \quad B] \sim \cdots \sim [I \quad X]$, entonces $X = A^{-1}B$.

Si A es más grande que 2×2, entonces la reducción por filas de $[A \quad B]$ es mucho más rápida que calcular A^{-1} y $A^{-1}B$.

13. Suponga que $AB = AC$, donde B y C son matrices $n \times p$ y A es invertible. Muestre que $B = C$. ¿Es esto cierto en general si A no es invertible?

14. Suponga $(B - C)D = 0$, donde B y C son matrices $m \times n$ y D es invertible. Muestre que $B = C$.

15. Suponga que A, B y C son matrices invertibles $n \times n$. Demuestre que ABC también es invertible construyendo una matriz D tal que $(ABC)D = I$ y $D(ABC) = I$.

16. Suponga que A y B son matrices $n \times n$, y que B y AB son invertibles. Muestre que A es invertible. [*Sugerencia:* Haga $C = AB$ y resuelva esta ecuación para A.]

17. Resuelva la ecuación $AB = BC$ para A, suponiendo que A, B y C son cuadradas y que B es invertible.

18. Suponga que P es invertible y $A = PBP^{-1}$. Despeje B en términos de A.

19. Si A, B y C son matrices invertibles $n \times n$, ¿la ecuación $C^{-1}(A + X)B^{-1} = I_n$ tiene alguna solución para X? Si es así, encuéntrela.

20. Suponga que A, B y X son matrices $n \times n$ con A, X, y $A - AX$ invertibles, y suponga que

$$(A - AX)^{-1} = X^{-1}B \tag{3}$$

a. Explique por qué B es invertible.

b. Resuelva (3) para X. Si es necesario invertir una matriz, explique por qué dicha matriz es invertible.

21. Explique por qué las columnas de una matriz A de $n \times n$ son linealmente independientes cuando A es invertible.

22. Explique por qué las columnas de una matriz A de $n \times n$ generan $\mathbb{R}^n$ cuando A es invertible. [*Sugerencia:* Revise el teorema 4 dado en la sección 1.4.]

23. Suponga que A es $n \times n$ y que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial. Explique por qué A tiene n columnas pivote y es equivalente por filas a I_n. Por el teorema 7, esto muestra que A debe ser invertible. (Este ejercicio y el 24 se citarán en la sección 2.3.)

24. Suponga que para una matriz A de $n \times n$ la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$. Explique ¿por qué A debe ser invertible? [*Sugerencia:* Considere si A es equivalente por filas a I_n.]

Los ejercicios 25 y 26 demuestran el teorema 4 para $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

25. Muestre que si $ad - bc = 0$, entonces la ecuación $A\mathbf{x} = \mathbf{0}$ tiene más de una solución. ¿Por qué implica esto que A no es invertible? [*Sugerencia:* Primero, considere $a = b = 0$. Después, si a y b no son ambos cero, considere el vector $\mathbf{x} = \begin{bmatrix} -b \\ a \end{bmatrix}$.]

26. Muestre que si $ad - bc \neq 0$, la fórmula para A^{-1} funciona.

Los ejercicios 27 y 28 demuestran casos especiales de los hechos acerca de matrices elementales establecidos en el recuadro que sigue al ejemplo 5. Aquí A es una matriz de 3×3 e $I = I_3$. (Una demostración general requeriría un poco más de notación.)

27. a. Use la ecuación (1) de la sección 2.1 para mostrar que la $\text{fila}_i(A) = \text{fila}_i(I) \cdot A$, para $i = 1, 2, 3$.

b. Muestre que si las filas 1 y 2 de A se intercambian, entonces el resultado puede escribirse como EA, donde E es una matriz elemental formada al intercambiar las filas 1 y 2 de I.

c. Muestre que si la fila 3 de A se multiplica por 5, entonces el resultado puede escribirse como EA, donde E se forma al multiplicar la fila 3 de I por 5.

28. Demuestre que si la fila 3 de A es reemplazada por $\text{fila}_3(A) - 4 \cdot \text{fila}_1(A)$, el resultado es EA, donde E se forma a partir de I mediante el reemplazo de $\text{fila}_3(I)$ por $\text{fila}_3(I) - 4 \cdot \text{fila}_1(I)$.

Encuentre los inversos de las matrices dadas en los ejercicios 29 a 32, si existen. Use el algoritmo presentado en esta sección.

29. $\begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix}$

30. $\begin{bmatrix} 5 & 10 \\ 4 & 7 \end{bmatrix}$

31. $\begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{bmatrix}$

32. $\begin{bmatrix} 1 & -2 & 1 \\ 4 & -7 & 3 \\ -2 & 6 & -4 \end{bmatrix}$

33. Use el algoritmo de esta sección para encontrar los inversos de $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ y $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. Sea A la matriz $n \times n$ correspondiente, y sea B su inverso. Estime la forma de B, y luego demuestre que $AB = I$ y $BA = I$.

34. Repita la estrategia del ejercicio 33 para obtener el inverso de $A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 0 & & 0 \\ 1 & 2 & 3 & & 0 \\ \vdots & & & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}$. Demuestre que su resultado es el correcto.

35. Sea $A = \begin{bmatrix} -2 & -7 & -9 \\ 2 & 5 & 6 \\ 1 & 3 & 4 \end{bmatrix}$. Encuentre la tercera columna de A^{-1} sin calcular las otras columnas.

36. **[M]** Sea $A = \begin{bmatrix} -25 & -9 & -27 \\ 546 & 180 & 537 \\ 154 & 50 & 149 \end{bmatrix}$. Encuentre la segunda y tercer columnas de A^{-1} sin calcular la primera columna.

37. Sea $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \end{bmatrix}$. Construya una matriz C de 2×3 (mediante prueba y error) usando sólo 1, -1 y 0 como entradas, de tal forma que $CA = I_2$. Calcule AC y observe que $AC \neq I_3$.

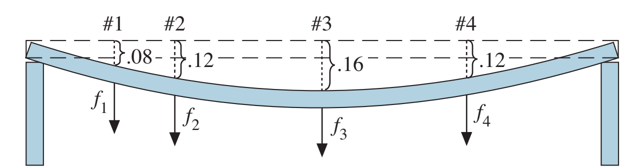
38. Sea $A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$. Construya una matriz D de 4×2 usando sólo 1 y 0 como entradas, de tal forma que $AD = I_2$. ¿Es posible que $CA = I_4$ para alguna matriz C de 4×2? ¿Por qué sí o por qué no?

39. Sea $D = \begin{bmatrix} .005 & .002 & .001 \\ .002 & .004 & .002 \\ .001 & .002 & .005 \end{bmatrix}$ una matriz de flexibilidad, con la flexibilidad medida en pulgadas por libra. Suponga que se aplican fuerzas de 30, 50 y 20 lb sobre los puntos 1, 2 y 3, respectivamente, en la figura 1 del ejemplo 3. Encuentre las deflexiones correspondientes.

40. **[M]** Encuentre la matriz de rigidez D^{-1} para la D del ejercicio 39. Enliste las fuerzas que se necesitan para producir una flexión de 0.04 pulgadas en el punto 3, con deflexión 0 en los otros puntos.

41. **[M]** Sea $D = \begin{bmatrix} .0040 & .0030 & .0010 & .0005 \\ .0030 & .0050 & .0030 & .0010 \\ .0010 & .0030 & .0050 & .0030 \\ .0005 & .0010 & .0030 & .0040 \end{bmatrix}$ una matriz de flexibilidad para una viga elástica con cuatro puntos en los cuales se aplican fuerzas. Las unidades son centímetros por newton de fuerza.

Las mediciones en los cuatro puntos muestran deflexiones de .08, .12, .16, y .12 cm. Determine las fuerzas presentes en los cuatro puntos.



Deflexión de una viga elástica para los ejercicios 41 y 42.

42. [M] Considere que D es como en el ejercicio 41, y determine las fuerzas que producen una deflexión de .24 cm en el segundo punto de la viga, con deflexión 0 en los otros tres puntos. ¿Qué relación hay entre la respuesta al problema y las entradas de D^{-1}? [*Sugerencia:* Primero conteste la pregunta para una deflexión de 1 cm en el segundo punto.]

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. a. $\det\begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix} = 3\cdot 6 - (-9)\cdot 2 = 18 + 18 = 36$. El determinante es diferente de cero, así que la matriz es invertible.

b. $\det\begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix} = 4\cdot 5 - (-9)\cdot 0 = 20 \neq 0$. La matriz es invertible.

c. $\det\begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix} = 6\cdot 6 - (-9)(-4) = 36 - 36 = 0$. La matriz no es invertible.

2.
$$[\,A \quad I\,] \sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ -1 & 5 & 6 & 0 & 1 & 0 \\ 5 & -4 & 5 & 0 & 0 & 1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ 0 & 3 & 5 & 1 & 1 & 0 \\ 0 & 6 & 10 & -5 & 0 & 1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ 0 & 3 & 5 & 1 & 1 & 0 \\ 0 & 0 & 0 & -7 & -2 & 1 \end{bmatrix}$$

Se ha obtenido una matriz de la forma $[B \quad D]$, donde B es cuadrada y tiene una fila de ceros. Las operaciones de fila adicionales no van a transformar B en I, así que el proceso se detiene. A no tiene un inverso.

2.3 CARACTERIZACIONES DE MATRICES INVERTIBLES

Esta sección proporciona un repaso de la mayor parte de los conceptos introducidos en el capítulo 1, en relación con sistemas de n ecuaciones lineales de n incógnitas y con matrices *cuadradas*. El resultado principal es el teorema 8.

TEOREMA 8 **El teorema de la matriz invertible**

Sea A una matriz cuadrada $n \times n$. Entonces, los siguientes enunciados son equivalentes. Esto es, para una A dada, los enunciados son o todos ciertos o todos falsos.

a. A es una matriz invertible.
b. A es equivalente por filas a la matriz identidad $n \times n$.
c. A tiene n posiciones pivote.
d. La ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.
e. Las columnas de A forman un conjunto linealmente independiente.
f. La transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno.
g. La ecuación $A\mathbf{x} = \mathbf{b}$ tiene por lo menos una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$.
h. Las columnas de A generan $\mathbb{R}^n$.
i. La transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^n$ sobre $\mathbb{R}^n$.
j. Existe una matriz C de $n \times$ n tal que $CA = I$.
k. Existe una matriz D de $n \times n$ tal que $AD = I$.
l. A^T es una matriz invertible.

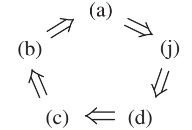


FIGURA 1

Primero, se necesita alguna notación. Si un enunciado (j) es cierto dado que algún enunciado (a) es cierto, se dice que (a) *implica* (j), y se escribe (a) $\Rightarrow$ (j). Se comenzará a establecer el "círculo" de implicaciones que muestra la figura 1. Si cualquiera de estos cinco enunciados es cierto, entonces también lo son los demás. Por último, se relacionarán los enunciados restantes del teorema con los enunciados incluidos en este círculo.

DEMOSTRACIÓN Si (a) es cierto, entonces A^{-1} funciona para C en (j), así (a) $\Rightarrow$ (j). Luego, (j) $\Rightarrow$ (d) por el ejercicio 23 de la sección 2.1. (Vuelva atrás y lea el ejercicio.) También, (d) $\Rightarrow$ (c) por el ejercicio 23 de la sección 2.2. Si A es cuadrada y tiene n posiciones pivote, entonces los pivotes deben estar sobre la diagonal principal, en cuyo caso, la forma escalonada reducida de A es I_n. Por lo tanto, (c) $\Rightarrow$ (b). También, (b) $\Rightarrow$ (a) por el teorema 7 de la sección 2.2. Esto completa el círculo de la figura 1.

(k)
(a) $\Leftarrow$ (g)

(g) $\Leftrightarrow$ (h) $\Leftrightarrow$ (i)

(d) $\Leftrightarrow$ (e) $\Leftrightarrow$ (f)

(a) $\Leftrightarrow$ (l)

Luego, (a) $\Rightarrow$ (k) porque A^{-1} funciona para D. También, (k) $\Rightarrow$ (g) por el ejercicio 24 de la sección 2.1, y (g) $\Rightarrow$ (a) por el ejercicio 24 de la sección 2.2. Así que (g) y (k) están conectados al círculo. Por otra parte, (g), (h) e (i) son equivalentes para cualquier matriz, por el teorema 4 de la sección 1.4 y el teorema 12(a) de la sección 1.9. Entonces, (h) e (i) también están conectados al círculo por medio de (g).

Como (d) está conectado al círculo, también lo están (e) y (f), porque (d), (e) y (f) son todos equivalentes para *cualquier* matriz A. (Vea la sección 1.7 y el teorema 12(b) de la sección 1.9.) Por último, (a) $\Rightarrow$ (l) por el teorema 6(c) de la sección 2.2, y (l) $\Rightarrow$ (a) por el mismo teorema intercambiando A y A^T. Esto completa la demostración. ■

Por el teorema 5 de la sección 2.2, el enunciado (g) del teorema 8 también podría escribirse como: "La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución *única* para toda $\mathbf{b}$ en $\mathbb{R}^n$". Desde luego que esta afirmación implica a (b) y, por lo tanto, implica que A es invertible.

El siguiente hecho es consecuencia del teorema 8 y del ejercicio 8 de la sección 2.2.

Sean A y B matrices cuadradas. Si $AB = I$, entonces A y B son invertibles, con $B = A^{-1}$ y $A = B^{-1}$.

El teorema de la matriz invertible divide al conjunto de todas las matrices $n \times n$ en dos clases excluyentes: las matrices invertibles (no singulares), y las matrices no invertibles (singulares). Cada enunciado del teorema describe una propiedad de toda matriz $n \times n$ invertible. La *negación* de un enunciado del teorema describe una propiedad de toda matriz singular $n \times n$. Por ejemplo, una matriz singular $n \times n$ *no* es equivalente por filas a I_n, *no* tiene n posiciones pivote, y tiene columnas linealmente *dependientes.* Las negaciones de los otros enunciados se consideran en los ejercicios.

EJEMPLO 1 Use el teorema de la matriz invertible para decidir si A es invertible:

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$$

Solución

$$A \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

Así que A tiene tres posiciones pivote y, por lo tanto, es invertible, por el teorema de la matriz invertible, enunciado (c).

SG Tabla expandida para la TMI 2 a 10 (Expanded Table for the IMT 2-10)

El poder del teorema de la matriz invertible radica en las conexiones que establece entre tantos conceptos importantes, tales como la independencia lineal de las columnas de una matriz A y la existencia de soluciones para ecuaciones de la forma $A\mathbf{x} = \mathbf{b}$. Sin embargo, debe subrayarse que el teorema de la matriz invertible *aplica solamente a matrices cuadradas.* Por ejemplo, si las columnas de una matriz de 4×3 son linealmente independientes, no puede usarse el teorema de la matriz invertible para obtener cualquier conclusión acerca de la existencia o no existencia de soluciones a ecuaciones de la forma $A\mathbf{x} = \mathbf{b}$.

Transformaciones lineales invertibles

Recuerde de la sección 2.1 que la multiplicación de matrices corresponde a la composición de transformaciones lineales. Cuando una matriz A es invertible, la ecuación $A^{-1}A\mathbf{x} = \mathbf{x}$ puede verse como un enunciado acerca de transformaciones lineales. Vea la figura 2.

Se dice que una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ es **invertible** si existe una función $S : \mathbb{R}^n \to \mathbb{R}^n$ tal que

$$S(T(\mathbf{x})) = \mathbf{x} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{R}^n \tag{1}$$

$$T(S(\mathbf{x})) = \mathbf{x} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{R}^n \tag{2}$$

El teorema siguiente muestra que si dicha S existe, es única y debe ser una transformación lineal. Se dice que S es el **inverso** de T y se escribe como T^{-1}.

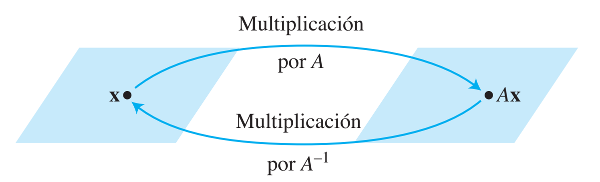


FIGURA 2 A^{-1}transforma a $A\mathbf{x}$ de nuevo en $\mathbf{x}$.

TEOREMA 9

Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal y sea A la matriz estándar para T. Entonces T es invertible si, y sólo si, A es una matriz invertible. En tal caso, la transformación lineal S dada por $S(\mathbf{x}) = A^{-1}\mathbf{x}$ es la función única que satisface (1) y (2).

DEMOSTRACIÓN Suponga que T es invertible. Entonces (2) muestra que T es sobre $\mathbb{R}^n$, porque si $\mathbf{b}$ está en $\mathbb{R}^n$ y $\mathbf{x} = S(\mathbf{b})$, entonces $T(\mathbf{x}) = T(S(\mathbf{b})) = \mathbf{b}$, así que toda $\mathbf{b}$ está en el rango de T. De manera que A es invertible, por el teorema de la matriz invertible, enunciado (i).

De manera inversa, suponga que A es invertible y sea $S(\mathbf{x}) = A^{-1}\mathbf{x}$. Entonces, S es una transformación lineal y S, desde luego, satisface (1) y (2). Por ejemplo;

$$S(T(\mathbf{x})) = S(A\mathbf{x}) = A^{-1}(A\mathbf{x}) = \mathbf{x}$$

Entonces T es invertible. La demostración de que S es única se describe de manera general en el ejercicio 39. ■

EJEMPLO 2 ¿Qué se puede decir acerca de una transformación lineal T uno a uno de $\mathbb{R}^n$ en $\mathbb{R}^n$?

Solución Las columnas de la matriz estándar A de T son linealmente independientes (por el teorema 12 de la sección 1.9). Así que A es invertible, por el teorema de la matriz invertible, y T mapea $\mathbb{R}^n$ *sobre* $\mathbb{R}^n$. También, T es invertible, por el teorema 9. ■

NOTAS NUMÉRICAS

En la práctica, puede encontrarse ocasionalmente una matriz “casi singular” o mal condicionada: una matriz invertible que puede convertirse en singular si algunas de sus entradas se cambian levemente. En este caso, la reducción por filas puede producir menos de n posiciones pivote, debido al error de redondeo. También, los errores de redondeo pueden algunas veces hacer que una matriz singular parezca ser invertible.

Algunos programas de matrices calculan un **número de condición** para una matriz cuadrada. Entre más grande sea el número de condición, más cerca estará la matriz de ser singular. El número de condición de la matriz identidad es 1. Una matriz singular tiene un número de condición infinito. En casos extremos, un programa de matrices podría no distinguir entre una matriz singular y una matriz mal condicionada.

Los ejercicios del 41 al 45 muestran que los cálculos de matrices pueden producir errores sustanciales cuando un número de condición es grande.

PROBLEMAS DE PRÁCTICA

1. Determine si $A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$ es invertible.
2. Suponga que para cierta matriz A de $n \times n$, el enunciado (g) del teorema de la matriz invertible *no* es verdadero. ¿Qué puede decirse acerca de las ecuaciones de la forma $A\mathbf{x} = \mathbf{b}$?
3. Suponga que A y B son matrices $n \times n$ y que la ecuación $AB\mathbf{x} = \mathbf{0}$ tiene una solución no trivial. ¿Qué puede decirse acerca de la matriz AB?

2.3 EJERCICIOS

A menos que se especifique lo contrario, suponga que en estos ejercicios todas las matrices son $n \times n$. En los ejercicios 1 a 10, determine cuáles de las matrices son invertibles. Use tan pocos cálculos como sea posible. Justifique sus respuestas.

1. $\begin{bmatrix} 5 & 7 \\ -3 & -6 \end{bmatrix}$

2. $\begin{bmatrix} -4 & 6 \\ 6 & -9 \end{bmatrix}$

3. $\begin{bmatrix} 5 & 0 & 0 \\ -3 & -7 & 0 \\ 8 & 5 & -1 \end{bmatrix}$

4. $\begin{bmatrix} -7 & 0 & 4 \\ 3 & 0 & -1 \\ 2 & 0 & 9 \end{bmatrix}$

5. $\begin{bmatrix} 0 & 3 & -5 \\ 1 & 0 & 2 \\ -4 & -9 & 7 \end{bmatrix}$

6. $\begin{bmatrix} 1 & -5 & -4 \\ 0 & 3 & 4 \\ -3 & 6 & 0 \end{bmatrix}$

7. $\begin{bmatrix} -1 & -3 & 0 & 1 \\ 3 & 5 & 8 & -3 \\ -2 & -6 & 3 & 2 \\ 0 & -1 & 2 & 1 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 3 & 7 & 4 \\ 0 & 5 & 9 & 6 \\ 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & 10 \end{bmatrix}$

9. **[M]** $\begin{bmatrix} 4 & 0 & -7 & -7 \\ -6 & 1 & 11 & 9 \\ 7 & -5 & 10 & 19 \\ -1 & 2 & 3 & -1 \end{bmatrix}$

10. **[M]** $\begin{bmatrix} 5 & 3 & 1 & 7 & 9 \\ 6 & 4 & 2 & 8 & -8 \\ 7 & 5 & 3 & 10 & 9 \\ 9 & 6 & 4 & -9 & -5 \\ 8 & 5 & 2 & 11 & 4 \end{bmatrix}$

En los ejercicios 11 y 12, todas las matrices son $n \times n$. Cada inciso de estos ejercicios es una *implicación* de la forma "si (enunciado 1), entonces (enunciado 2)". Califique una implicación como verdadera si (enunciado 2) es verdadero *siempre* que (enunciado 1) sea cierto. Una implicación es falsa si existe una instancia en la que (enunciado 2) es falso pero (enunciado 1) es verdadero. Justifique sus respuestas.

11. a. Si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial, entonces A es equivalente por filas a la matriz identidad de $n \times n$.
    b. Si las columnas de A generan $\mathbb{R}^n$, entonces las columnas son linealmente independientes.
    c. Si A es una matriz de $n \times n$, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$.
    d. Si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene una solución no trivial, entonces A tiene menos de n posiciones pivote.
    e. Si A^T no es invertible, entonces A no es invertible.
12. a. Si existe una matriz D de $n \times n$ tal que $AD = I$, entonces también existe una matriz C de $n \times n$ tal que $CA = I$.
    b. Si las columnas de A son linealmente independientes, entonces las columnas de A generan $\mathbb{R}^n$.
    c. Si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$, entonces la solución es única para toda $\mathbf{b}$.
    d. Si la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es una función de $\mathbb{R}^n$ en $\mathbb{R}^n$, entonces A tiene n posiciones pivote.
    e. Si existe una $\mathbf{b}$ en $\mathbb{R}^n$ tal que la ecuación $A\mathbf{x} = \mathbf{b}$ sea inconsistente, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no es uno a uno.
13. Una **matriz triangular superior** de $m \times n$ es aquella cuyas entradas *abajo* de la diagonal principal son ceros (como en el ejercicio 8). ¿Cuándo es invertible una matriz triangular superior cuadrada? Justifique su respuesta.
14. Una **matriz triangular inferior** de $m \times n$ es aquella cuyas entradas *arriba* de la diagonal principal son ceros (como en el ejercicio 3). ¿Cuándo es invertible una matriz triangular inferior cuadrada? Justifique su respuesta.
15. ¿Puede ser invertible una matriz cuadrada con dos columnas idénticas? ¿Por qué sí o por qué no?

16. ¿Es posible que una matriz de 5×5 sea invertible cuando sus columnas no generan $\mathbb{R}^n$. ¿Por qué sí o por qué no?

17. Si A es invertible, las columnas de A^{-1} son linealmente independientes. Explique por qué.

18. Si C es de 6×6 y la ecuación $C\mathbf{x} = \mathbf{v}$ es consistente para toda $\mathbf{v}$ en $\mathbb{R}^6$, ¿es posible que la ecuación $C\mathbf{x} = \mathbf{v}$ tenga más de una solución para alguna $\mathbf{v}$? ¿Por qué sí o por qué no?

19. Si las columnas de una matriz de 7×7 son linealmente independientes, ¿qué puede decirse acerca de las soluciones de $D\mathbf{x} = \mathbf{b}$? ¿Por qué?

20. Si las matrices de $n \times n$ E y F tienen la propiedad de que $EF = I$, entonces E y F conmutan. Explique por qué.

21. Si la ecuación $G\mathbf{x} = \mathbf{y}$ tiene más de una solución para alguna $\mathbf{y}$ en $\mathbb{R}^n$, ¿las columnas de G generan a $\mathbb{R}^n$? ¿Por qué sí o por qué no?

22. Si la ecuación $H\mathbf{x} = \mathbf{c}$ es inconsistente para alguna $\mathbf{c}$ en $\mathbb{R}^n$, ¿qué puede decirse acerca de la ecuación $H\mathbf{x} = \mathbf{0}$? ¿Por qué?

23. Si una matriz K de $n \times n$ no puede reducirse por filas a I_n, ¿qué puede decirse de las columnas de K? ¿Por qué?

24. Si L es $n \times n$ y la ecuación $L\mathbf{x} = \mathbf{0}$ tiene la solución trivial, ¿las columnas de L generan a $\mathbb{R}^n$? ¿Por qué?

25. Verifique el enunciado del recuadro que sigue al ejemplo 1.

26. Explique por qué las columnas de A^2 generan $\mathbb{R}^n$ siempre que las columnas de A son linealmente independientes.

27. Demuestre que si AB es invertible, también lo es A. No puede usarse el teorema 6(b), porque no es posible *suponer* que A y B son invertibles. [*Sugerencia:* Existe una matriz W tal que $ABW = I$. ¿Por qué?]

28. Demuestre que si AB es invertible, también B lo es.

29. Si A es una matriz $n \times n$ y la ecuación $A\mathbf{x} = \mathbf{b}$ tiene más de una solución para alguna $\mathbf{b}$, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no es uno a uno. ¿Qué otra cosa puede decirse acerca de esta transformación? Justifique su respuesta.

30. Si A es una matriz de $n \times n$ y la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno, ¿qué otra cosa puede decirse acerca de esta transformación? Justifique su respuesta.

31. Suponga que A es una matriz $n \times n$ con la propiedad de que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para cada $\mathbf{b}$ en $\mathbb{R}^n$. Sin utilizar los teoremas 5 u 8, explique por qué cada ecuación $A\mathbf{x} = \mathbf{b}$ tiene, de hecho, exactamente una solución.

32. Suponga que A es una matriz $n \times n$ con la propiedad de que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial. Sin utilizar el teorema de la matriz invertible, explique directamente por qué la ecuación $A\mathbf{x} = \mathbf{b}$ debe tener una solución para cada $\mathbf{b}$ en $\mathbb{R}^n$.

En los ejercicios 33 y 34, T es una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$. Demuestre que T es invertible y encuentre una fórmula para T^{-1}.

33. $T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2)$

34. $T(x_1, x_2) = (6x_1 - 8x_2, -5x_1 + 7x_2)$

35. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal. Explique por qué T es tanto uno a uno como sobre $\mathbb{R}^n$. Use las ecuaciones (1) y (2). Luego, dé una segunda explicación usando uno o más teoremas.

36. Sea T una transformación lineal que mapea $\mathbb{R}^n$ sobre $\mathbb{R}^n$. Muestre que T^{-1} existe y mapea $\mathbb{R}^n$ en $\mathbb{R}^n$. ¿T^{-1} es también uno a uno?

37. Suponga que T y U son transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^n$ tales que $T(U(\mathbf{x})) = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$. ¿Es cierto que $U(T(\mathbf{x})) = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$ ¿Por qué sí o por qué no?

38. Suponga una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ con la propiedad de que $T(\mathbf{u}) = T(\mathbf{v})$ para algún par de vectores distintos $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$. ¿Puede T mapear $\mathbb{R}^n$ sobre $\mathbb{R}^n$? ¿Por qué sí o por qué no?

39. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal invertible, y sean S y U funciones de $\mathbb{R}^n$ en $\mathbb{R}^n$ tales que $S(T(\mathbf{x})) = \mathbf{x}$ y $U(T(\mathbf{x})) = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$. Muestre que $U(\mathbf{v}) = S(\mathbf{v})$ para toda $\mathbf{v}$ en $\mathbb{R}^n$. Esto demostrará que T tiene un inverso único, como se afirma en el teorema 9. [*Sugerencia:* Dada cualquier $\mathbf{v}$ en $\mathbb{R}^n$, se puede escribir $\mathbf{v} = T(\mathbf{x})$ para alguna $\mathbf{x}$. ¿Por qué? Calcule $S(\mathbf{v})$ y $U(\mathbf{v})$.]

40. Suponga que T y S satisfacen las ecuaciones de invertibilidad (1) y (2), donde T es una transformación lineal. Muestre directamente que S es una transformación lineal. [*Sugerencia:* Dadas $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, sea $\mathbf{x} = S(\mathbf{u})$, $\mathbf{y} = S(\mathbf{v})$. Entonces $T(\mathbf{x}) = \mathbf{u}$, $T(\mathbf{y}) = \mathbf{v}$. ¿Por qué? Aplique S a ambos miembros de la ecuación $T(\mathbf{x}) + T(\mathbf{y}) = T(\mathbf{x} + \mathbf{y})$. También, considere $T(c\mathbf{x}) = cT(\mathbf{x})$.]

41. **[M]** Suponga que un experimento conduce al siguiente sistema de ecuaciones:

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.249 \\ 1.6x_1 + 1.1x_2 &= 6.843 \end{aligned} \qquad (3)$$

a. Resuelva el sistema (3), y después resuelva el sistema (4) que se muestra a continuación, en el cual los datos a la derecha se han redondeado a dos decimales. En cada caso. encuentre la solución *exacta.*

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.25 \\ 1.6x_1 + 1.1x_2 &= 6.84 \end{aligned} \qquad (4)$$

b. Las entradas de (4) difieren de las de (3) en menos de 0.05%. Encuentre el porcentaje de error cuando se utiliza la solución de (4) como una aproximación a la solución de (3).

c. Use un programa de matrices para producir el número de condición de la matriz de coeficientes de (3).

Los ejercicios 42, 43 y 44 muestran cómo utilizar el número de condición de una matriz A para estimar la exactitud de una solución calculada de $A\mathbf{x} = \mathbf{b}$. Si las entradas de A y $\mathbf{b}$ son exactas hasta más o menos r dígitos significativos, y si el número de condición de A es aproximadamente 10^k (siendo k un entero positivo), entonces la solución calculada de $A\mathbf{x} = \mathbf{b}$ debería ser exacta hasta al menos $r - k$ dígitos significativos.

42. **[M]** Encuentre el número de condición de la matriz A en el ejercicio 9. Construya un vector $\mathbf{x}$ al azar en $\mathbb{R}^4$ y calcule $\mathbf{b} = A\mathbf{x}$. Después use un programa de matrices para calcular la solución $\mathbf{x}_1$ de $A\mathbf{x} = \mathbf{b}$. ¿Hasta cuántos dígitos coinciden $\mathbf{x}$ y $\mathbf{x}_1$? Encuentre el número de dígitos que el programa de matrices almacena con precisión, e informe acerca de cuántos dígitos de exactitud se pierden cuando se usa $\mathbf{x}_1$ en lugar de la solución exacta $\mathbf{x}$.

43. **[M]** Repita el ejercicio 42 para la matriz del ejercicio 10.

44. **[M]** Resuelva la ecuación $A\mathbf{x} = \mathbf{b}$ para obtener una $\mathbf{b}$ que sirva para encontrar la última columna del inverso de la *matriz de Hilbert de quinto orden*

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{bmatrix}$$

¿Cuántos dígitos de cada entrada de $\mathbf{x}$ puede esperarse que sean correctos? Explique su respuesta. [*Nota:* La solución exacta es (630, −12 600, 56 700, −88 200, 44 100).]

45. **[M]** Algunos programas de matrices, como MATLAB, tienen una orden para crear matrices de Hilbert de varios tamaños. Si es posible, use una orden inversa para calcular el inverso de una matriz de Hilbert A de orden doce o mayor. Calcule AA^{-1}. Informe acerca de sus descubrimientos.

SG Dominio: Revisión y reflexión 2 a 13 (Mastering: Reviewing and Reflecting 2-13)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Es evidente que las columnas de A son linealmente dependientes porque las columnas 2 y 3 son múltiplos de la columna 1. Por lo tanto, A no puede ser invertible, por el teorema de la matriz invertible.
2. Si el enunciado (g) *no* es cierto, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es inconsistente para, por lo menos, una $\mathbf{b}$ en $\mathbb{R}^n$.
3. Aplique el teorema de la matriz invertible a la matriz AB en lugar de A. Entonces el enunciado (d) se convierte en: $AB\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial. Esto no es cierto. Por lo tanto, AB no es invertible.

2.4 MATRICES PARTIDAS

Una característica clave del trabajo con matrices realizado hasta aquí ha sido la capacidad para considerar a una matriz A como una lista de vectores columna en lugar de, simplemente, un arreglo rectangular de números. Este punto de vista ha resultado tan útil que sería deseable considerar otras **particiones** de A, indicadas mediante líneas divisorias horizontales y verticales, como en el ejemplo 1 que se presenta a continuación. Las matrices partidas aparecen con frecuencia en las aplicaciones modernas del álgebra lineal porque la notación simplifica muchos análisis y resalta la estructura esencial de los cálculos matriciales, como se mostró en el ejemplo introductorio de este capítulo acerca del diseño de aviones. Esta sección proporciona una oportunidad para revisar el álgebra matricial y usar el teorema de la matriz invertible.

EJEMPLO 1 La matriz

$$A = \left[\begin{array}{ccc|cc|c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

también puede escribirse como la **matriz partida** (o en **bloques**) de 2 × 3

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

cuyas entradas son los *bloques* (o *submatrices*)

$$A_{11} = \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 1 & 7 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} -4 \end{bmatrix}$$

EJEMPLO 2 Cuando una matriz A aparece en un modelo matemático de un sistema físico, tal como en una red eléctrica, un sistema de transporte, o una gran compañía, puede resultar natural considerar a A como una matriz partida. Por ejemplo, si un tablero de circuitos de microcomputadora consiste, principalmente, en tres microcircuitos VLSI (del inglés *very large-scale integrated:* integrados a escala muy grande), entonces la matriz para el tablero de circuitos podría tener la forma general

$$A = \left[\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ \hline A_{31} & A_{32} & A_{33} \end{array}\right]$$

Las submatrices sobre la "diagonal" de A —a saber, A_{11}, A_{22} y A_{33}— se refieren a los tres circuitos VLSI, mientras que las otras submatrices dependen de las interconexiones que haya entre esos microcircuitos.

Suma y multiplicación escalares

Si las matrices A y B son del mismo tamaño y están partidas exactamente en la misma forma, entonces es natural efectuar una partición similar de la suma ordinaria matricial $A + B$. En este caso, cada bloque de $A + B$ es la suma (matricial) de los bloques correspondientes de A y B. La multiplicación por un escalar de una matriz partida también se calcula bloque por bloque.

Multiplicación de matrices partidas

Las matrices partidas se pueden multiplicar mediante la regla acostumbrada fila-columna como si las entradas del bloque fueran escalares, siempre y cuando, para un producto AB, la partición por columnas de A equivalga a la partición por filas de B.

EJEMPLO 3 Sean

$$A = \left[\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \left[\begin{array}{cc} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

Las cinco columnas de A están partidas en un conjunto de tres columnas y luego en uno de dos columnas. Las cinco filas de B están partidas de igual manera —en un conjunto de tres filas y luego en uno de dos filas. Se dice que las particiones de A y B están **conformadas** para **multiplicación de bloques**. Es posible mostrar que el producto común AB puede escribirse como

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} = \left[\begin{array}{rr} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right]$$

Es importante escribir cada producto menor de la expresión para AB con la submatriz de A a la izquierda, dado que la multiplicación de matrices no es conmutativa. Por ejemplo,

$$A_{11}B_1 = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 5 & -2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix}$$

$$A_{12}B_2 = \begin{bmatrix} 0 & -4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix}$$

Por lo tanto, el bloque superior es

$$A_{11}B_1 + A_{12}B_2 = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix} + \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix} = \begin{bmatrix} -5 & 4 \\ -6 & 2 \end{bmatrix}$$

La regla fila-columna para la multiplicación de matrices en bloques proporciona la manera más general de considerar un producto de dos matrices. Cada una de las siguientes formas de ver un producto ya se ha descrito usando particiones sencillas de matrices: (1) la definición de $A\mathbf{x}$ usando las columnas de A, (2) la definición de columna de AB, (3) la regla fila-columna para calcular AB, y (4) las filas de AB como productos de las filas de A y la matriz B. Una quinta manera de ver AB, también usando particiones, se dará posteriormente en el teorema 10.

Los cálculos del siguiente ejemplo preparan el camino para el teorema 10. Aquí, $\text{col}_k(A)$ es la k-ésima columna de A, y $\text{fila}_k(B)$ es la k-ésima fila de B.

EJEMPLO 4 Sean $A = \begin{bmatrix} -3 & 1 & 2 \\ 1 & -4 & 5 \end{bmatrix}$ y $B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$. Verifique que

$$AB = \text{col}_1(A)\,\text{fila}_1(B) + \text{col}_2(A)\,\text{fila}_2(B) + \text{col}_3(A)\,\text{fila}_3(B)$$

Solución Cada uno de los términos anteriores es un *producto externo*. (Vea los ejercicios 27 y 28 de la sección 2.1.) Por la regla fila-columna para calcular un producto matricial,

$$\text{col}_1(A)\,\text{fila}_1(B) = \begin{bmatrix} -3 \\ 1 \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} -3a & -3b \\ a & b \end{bmatrix}$$

$$\text{col}_2(A)\,\text{fila}_2(B) = \begin{bmatrix} 1 \\ -4 \end{bmatrix} \begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} c & d \\ -4c & -4d \end{bmatrix}$$

$$\text{col}_3(A)\,\text{fila}_3(B) = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \begin{bmatrix} e & f \end{bmatrix} = \begin{bmatrix} 2e & 2f \\ 5e & 5f \end{bmatrix}$$

Entonces

$$\sum_{k=1}^{3} \text{col}_k(A)\,\text{fila}_k(B) = \begin{bmatrix} -3a + c + 2e & -3b + d + 2f \\ a - 4c + 5e & b - 4d + 5f \end{bmatrix}$$

Resulta evidente que esta matriz es AB. Observe que la entrada (1, 1) de AB es la suma de las entradas (1, 1) de los tres productos externos, la entrada (1, 2) de AB es la suma de las entradas (1, 2) de los tres productos externos, y así sucesivamente. ■

TEOREMA 10 **Ampliación columna-fila de AB**

Si A es $m \times n$ y B es $n \times p$, entonces

$$AB = [\,\text{col}_1(A) \quad \text{col}_2(A) \quad \cdots \quad \text{col}_n(A)\,] \begin{bmatrix} \text{fila}_1(B) \\ \text{fila}_2(B) \\ \vdots \\ \text{fila}_n(B) \end{bmatrix} \tag{1}$$

$$= \text{col}_1(A)\,\text{fila}_1(B) + \cdots + \text{col}_n(A)\,\text{fila}_n(B)$$

DEMOSTRACIÓN Para cada índice de fila i e índice de columna j, la entrada (i, j) en $\text{col}_k(A)\,\text{fila}_k(B)$ es el producto de a_{ik} de $\text{col}_k(A)$ y b_{kj} de $\text{fila}_k(B)$. Por lo tanto, la entrada (i, j) de la suma que muestra (1) es

$$\underset{(k=1)}{a_{i1}b_{1j}} + \underset{(k=2)}{a_{i2}b_{2j}} + \cdots + \underset{(k=n)}{a_{in}b_{nj}}$$

Esta suma también es la entrada (i, j) de AB, por la regla fila-columna. ■

Inversos de matrices partidas

El siguiente ejemplo ilustra los cálculos relacionados con inversos y matrices partidas.

EJEMPLO 5 Una matriz de la forma

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

se dice que es *triangular superior en bloques*. Suponga que A_{11} es $p \times p$, A_{22} $q \times q$, y A invertible. Encuentre una fórmula para A^{-1}.

Solución Denote A^{-1} mediante B, y efectúe una partición de B para que

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_q \end{bmatrix} \tag{2}$$

Esta ecuación matricial proporciona cuatro ecuaciones que conducen a las submatrices desconocidas $B_{11}, \ldots, B_{22}$. Calcule el producto a la izquierda de (2), e iguale

cada entrada con el bloque correspondiente en la matriz identidad a la derecha. Esto es, establezca

$$A_{11}B_{11} + A_{12}B_{21} = I_p \quad (3)$$

$$A_{11}B_{12} + A_{12}B_{22} = 0 \quad (4)$$

$$A_{22}B_{21} = 0 \quad (5)$$

$$A_{22}B_{22} = I_q \quad (6)$$

Por sí misma, (6) no establece que A_{22} sea invertible, porque todavía no se sabe que $B_{22}A_{22} = I_q$. Pero, al aplicar el teorema de la matriz invertible y el hecho de que A_{22} es cuadrada, *puede* concluirse que A_{22} es invertible y $B_{22} = A_{22}^{-1}$. Ahora se puede usar (5) para encontrar

$$B_{21} = A_{22}^{-1}0 = 0$$

así que (3) se simplifica a

$$A_{11}B_{11} + 0 = I_p$$

Esto demuestra que A_{11} es invertible y $B_{11} = A_{11}^{-1}$. Por último, de la expresión (4),

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \quad \text{y} \quad B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

Así que

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$$

Una **matriz diagonal en bloques** es una matriz partida con bloques de ceros fuera de la diagonal (de bloques) principal. Una matriz de este tipo es invertible si, y sólo si, cada bloque sobre la diagonal es invertible. Vea los ejercicios 13 y 14.

NOTAS NUMÉRICAS

1. Cuando las matrices son demasiado grandes para caber en la memoria de alta velocidad de una computadora, partirlas permite a la computadora trabajar solamente con dos o tres submatrices a la vez. Por ejemplo, en trabajos recientes sobre programación lineal, un equipo de investigación simplificó un problema al partir la matriz en 837 filas y 51 columnas. La resolución del problema tardó aproximadamente cuatro minutos en una supercomputadora Cray.[1]
2. Algunas computadoras de alta velocidad, en particular aquellas con arquitectura de conducción vectorial, realizan cálculos matriciales con mayor eficiencia cuando los algoritmos usan matrices partidas.[2]
3. Los programas de cómputo profesionales para álgebra lineal numérica de alto desempeño, LAPACK, hacen un uso intensivo de cálculos de matrices partidas.

[1]El tiempo de resolución no parece muy impresionante, hasta saber que cada bloque de las 51 columnas contenía, aproximadamente, 250,000 columnas individuales. ¡El problema original tenía 837 ecuaciones y más de 12,750,000 variables! Casi 100 millones de las más de 10 mil millones de entradas eran diferentes de cero. Vea Robert E. Bixby *et al.*, "Very Large-Scale Linear Programming: A Case Study in Combining Interior Point and Simplex Methods", *Operations Research*, 40, núm. 5 (1992): págs. 885-897.

[2]La importancia de los algoritmos de matrices en bloque para cálculos de computadora se describe en *Matrix Computations*, 3a. ed., por Gene H. Golub y Charles F. van Loan (Baltimore: Johns Hopkins University Press, 1996).

Los ejercicios siguientes permiten practicar el álgebra matricial, e ilustran cálculos típicos que pueden encontrarse durante las aplicaciones.

PROBLEMAS DE PRÁCTICA

1. Muestre que $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ es invertible y encuentre su inversa.

2. Calcule X^TX, cuando X está partida como $\begin{bmatrix} X_1 & X_2 \end{bmatrix}$.

2.4 EJERCICIOS

En los ejercicios 1 a 9, suponga que las matrices están partidas de manera adecuada para la multiplicación por bloques. Encuentre los productos mostrados en los ejercicios 1 a 4.

1. $\begin{bmatrix} I & 0 \\ E & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

2. $\begin{bmatrix} E & 0 \\ 0 & F \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

3. $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$

4. $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

En los ejercicios 5 a 8, encuentre fórmulas para X, Y y Z en términos de A, B y C, y justifique sus cálculos. Para producir una fórmula, en algunos casos, puede ser necesario formular suposiciones acerca del tamaño de una matriz. [*Sugerencia:* Calcule el producto de la izquierda e iguálelo al miembro del lado derecho.]

5. $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ X & Y \end{bmatrix} = \begin{bmatrix} 0 & I \\ Z & 0 \end{bmatrix}$

6. $\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}\begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

7. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

8. $\begin{bmatrix} A & B \\ 0 & I \end{bmatrix}\begin{bmatrix} X & Y & Z \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$

9. Suponga que A_{11} es una matriz invertible. Encuentre matrices X y Y tales que el producto mostrado a continuación tenga la forma indicada. También, calcule B_{22}. [*Sugerencia:* Calcule el producto de la izquierda e iguálelo al miembro del lado derecho.]

$$\begin{bmatrix} I & 0 & 0 \\ X & I & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \\ 0 & B_{32} \end{bmatrix}$$

10. El inverso de $\begin{bmatrix} I & 0 & 0 \\ C & I & 0 \\ A & B & I \end{bmatrix}$ es $\begin{bmatrix} I & 0 & 0 \\ Z & I & 0 \\ X & Y & I \end{bmatrix}$. Encuentre X, Y y Z.

En los ejercicios 11 y 12, señale cada afirmación como verdadera o falsa. Justifique sus respuestas.

11. a. Si $A = [A_1 \quad A_2]$ y $B = [B_1 \quad B_2]$, teniendo A_1 y A_2 el mismo tamaño que B_1 y B_2, respectivamente, entonces $A + B = [A_1 + B_1 \quad A_2 + B_2]$.

b. Si $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ y $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, entonces las particiones de A y B están conformadas para multiplicación por bloques.

12. a. La definición del producto matriz-vector $A\mathbf{x}$ es un caso especial de la multiplicación por bloques.

b. Si A_1, A_2, B_1 y B_2 son matrices de $n \times n$, $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$, y $B = [B_1 \quad B_2]$, entonces el producto BA está definido, pero AB no.

13. Sea $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$, donde B y C son cuadradas. Demuestre que A es invertible si, y sólo si, tanto B como C son invertibles.

14. Muestre que la matriz triangular superior en bloque A presentada en el ejemplo 5 es invertible si, y sólo si, tanto A_{11} como A_{22} son invertibles. [*Sugerencia:* Si A_{11} y A_{22} son invertibles, la fórmula para A^{-1} dada en el ejemplo 5 funciona realmente como el inverso de A.] Este hecho acerca de A es una parte importante de varios algoritmos de computadora que estiman valores propios de matrices. Los valores propios se analizan en el capítulo 5.

15. Suponga que A_{11} es invertible. Encuentre X y Y tales que

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} \quad (7)$$

Donde $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$. La matriz S es el **complemento de Schur** de A_{11}. De igual modo, si A_{22} es invertible, la matriz $A_{11} - A_{12}A_{22}^{-1}A_{21}$ se denomina complemento de Schur de A_{22}. Tales expresiones son comunes en la teoría de ingeniería de sistemas y en otras áreas.

16. Suponga que la matriz de bloques A ubicada en el miembro izquierdo de (7) y A_{11} son invertibles. Demuestre que el complemento de Schur S de A_{11} es invertible. [*Indicación:* Los factores externos localizados en el miembro derecho de (7) siempre son invertibles. Verifique esto.] Cuando A y A_{11} son ambos invertibles, (7) conduce a una fórmula para A^{-1}, utilizando S^{-1}, A_{11}^{-1}, y las otras entradas de A.

17. Cuando se lanza una sonda al espacio profundo, puede ser necesario efectuar correcciones para colocarla en una trayectoria calculada con precisión. La telemetría radial proporciona una serie de vectores, $\mathbf{x}_1, \dots, \mathbf{x}_k$, que dan información en diversos momentos acerca de la diferencia entre la posición de la sonda y su trayectoria planeada. Sea X_k la matriz $[\mathbf{x}_1 \cdots \mathbf{x}_k]$. La matriz $G_k = X_k X_k^T$ se calcula conforme se analizan los datos del radar. Cuando llega $\mathbf{x}_{k+1}$, se debe calcular una nueva G_{k+1}. Dado que los vectores de datos llegan a alta velocidad, la carga computacional podría ser severa. Sin embargo, la multiplicación de matrices proporciona una ayuda muy grande. Determine los desarrollos columna-fila de G_k y G_{k+1}, y describa lo que se debe calcular para *actualizar* G_k y formar G_{k+1}.

La sonda Galileo fue lanzada el 18 de octubre de 1989, y llegó cerca de Júpiter los primeros días de diciembre de 1995.

18. Sea X una matriz de datos de $m \times n$ tal que X^TX es invertible, y sea $M = I_m - X(X^TX)^{-1}X^T$. Añada una columna $\mathbf{x}_0$ a los datos y forme

$$W = [X \quad \mathbf{x}_0].$$

Calcule W^TW. La entrada (1, 1) es X^TX. Muestre que el complemento de Schur (ejercicio 15) de X^TX puede escribirse en la forma $\mathbf{x}_0{}^T M \mathbf{x}_0$. Se puede demostrar que la cantidad $(\mathbf{x}_0{}^T M \mathbf{x}_0)^{-1}$ es la entrada (2, 2) de $(W^TW)^{-1}$. Esta entrada tiene una interpretación estadística útil, bajo las hipótesis apropiadas.

En el estudio de ingeniería de control de sistemas físicos, un conjunto estándar de ecuaciones diferenciales se convierte en el siguiente sistema de ecuaciones lineales por medio de transformadas de Laplace:

$$\begin{bmatrix} A - sI_n & B \\ C & I_m \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} \tag{8}$$

donde A es de $n \times n$, B de $n \times m$, C de $m \times n$, y s una variable. El vector $\mathbf{u}$ en $\mathbb{R}^m$ es la "entrada" del sistema, $\mathbf{y}$ en $\mathbb{R}^m$ es la "salida" del sistema, y $\mathbf{x}$ en $\mathbb{R}^n$ es el vector de "estado". (De hecho, los vectores $\mathbf{x}$, $\mathbf{u}$ e $\mathbf{y}$ son funciones de s, pero este hecho se omite porque no afecta los cálculos algebraicos de los ejercicios 19 y 20.)

19. Suponga que $A - sI_n$ es invertible y vea a (8) como un sistema de dos ecuaciones matriciales. Resuelva la ecuación superior para $\mathbf{x}$ y sustitúyala en la ecuación inferior. El resultado es una ecuación de la forma $W(s)\mathbf{u} = \mathbf{y}$, donde $W(s)$ es una matriz que depende de s. $W(s)$ se denomina *función de transferencia* del sistema porque transforma la entrada $\mathbf{u}$ en la salida $\mathbf{y}$. Encuentre $W(s)$ y describa cómo está relacionada con el *sistema de matriz* partida del miembro izquierdo de (8). Vea el ejercicio 15.

20. Suponga que la función de transferencia $W(s)$ del ejercicio 19 es invertible para alguna s. Puede mostrarse que la función de transferencia inversa $W(s)^{-1}$, la cual transforma salidas en entradas, es el complemento de Schur de $A - BC - sI_n$ para la matriz que se presenta a continuación. Encuentre este complemento de Schur. Vea el ejercicio 15.

$$\begin{bmatrix} A - BC - sI_n & B \\ -C & I_m \end{bmatrix}$$

21. a. Verifique que $A^2 = I$ cuando $A = \begin{bmatrix} 1 & 0 \\ 3 & -1 \end{bmatrix}$.

b. Use matrices partidas para demostrar que $M^2 = I$ cuando

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -3 & 1 \end{bmatrix}.$$

22. Generalice la idea del ejercicio 21(a) [no del 21(b)] al construir una matriz de 5×5, $M = \begin{bmatrix} A & 0 \\ C & D \end{bmatrix}$ tal que $M^2 = I$. Convierta a C en una matriz de 2×3 distinta de cero. Muestre que su estructura funciona.

23. Use matrices partidas para demostrar por inducción que el producto de dos matrices triangulares inferiores es también triangular inferior. [*Sugerencia:* Una matriz A_1 de $(k + 1) \times (k + 1)$ puede escribirse en la forma presentada a continuación, donde a es un escalar, $\mathbf{v}$ está en $\mathbb{R}^k$, y A es una matriz triangular inferior de $k \times k$. [Vea la *guía de estudio* (*Study Guide*) para obtener ayuda con la inducción.]

$$A_1 = \begin{bmatrix} a & \mathbf{0}^T \\ \mathbf{v} & A \end{bmatrix}.$$

SG El principio de inducción 2 a 20 (The Principle of Induction 2-20)

24. Use matrices partidas para demostrar por inducción que para $n = 2, 3, \ldots$, la matriz A de $n \times n$ presentada a continuación es invertible y que B es su inverso.

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & & 0 \\ 1 & 1 & 1 & & 0 \\ \vdots & & & \ddots & \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & & 0 \\ 0 & -1 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & & \ldots & -1 & 1 \end{bmatrix}$$

Para el paso de inducción, suponga que A y B son matrices de $(k + 1) \times (k + 1)$, y parta A y B de una manera similar a la desplegada en el ejercicio 23.

25. Sin utilizar reducción por filas, encuentre el inverso de

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 5 & 6 \end{bmatrix}$$

26. **[M]** Para las operaciones de bloque, podría ser necesario introducir o recurrir a submatrices de una matriz grande. Describa las funciones o comandos de un programa de matrices que realizan las siguientes tareas. Suponga que A es una matriz de 20×30.

a. Desplegar la submatriz de A desde las filas 15 a 20 y las columnas 5 a 10.

b. Insertar en A una matriz B de 5×10, comenzando en la fila 10 y la columna 20.

c. Crear una matriz de 50×50 de la forma $B = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}$.

[*Nota:* Podría no ser necesario especificar los bloques de sólo ceros en B.]

27. **[M]** Suponga que debido a restricciones de memoria o tamaño su programa de matrices no puede trabajar con matrices de más de 32 filas y 32 columnas, y suponga que algún proyecto requiere las matrices A y B de 50×50. Describa los comandos u operaciones de su programa para matrices que realizan las siguientes tareas.

a. Calcular $A + B$.

b. Calcular AB.

c. Resolver $A\mathbf{x} = \mathbf{b}$ para algún vector $\mathbf{b}$ en $\mathbb{R}^{50}$, suponiendo que A se puede partir en una matriz de bloque de 2×2 $[A_{ij}]$, siendo A_{11} una matriz invertible de 20×20, A_{22} una matriz invertible de 30×30, y A_{12} una matriz cero. [*Sugerencia:* Describa sistemas apropiados más pequeños que puedan resolverse sin usar inversos de matrices.]

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Si $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ es invertible, su inverso tiene la forma $\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$. Se calcula

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix} = \begin{bmatrix} W & X \\ AW + Y & AX + Z \end{bmatrix}$$

Así que W, X, Y, Z deben satisfacer $W = I$, $X = 0$, $AW + Y = 0$, y $AX + Z = I$. Se sigue que $Y = -A$ y $Z = I$. Por lo tanto,

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

El producto en el orden inverso es también la identidad, de modo que la matriz de bloque es invertible, y su inverso es $\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix}$. (También se podría recurrir al teorema de la matriz invertible.)

2. $X^TX = \begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^TX_1 & X_1^TX_2 \\ X_2^TX_1 & X_2^TX_2 \end{bmatrix}$. Las particiones de X^T y X son conformadas de manera automática para la multiplicación porque las columnas de X^T son las filas de X. Esta partición de X^TX se usa en varios algoritmos de computadora para efectuar cálculos de matrices.

2.5 FACTORIZACIONES DE MATRICES

La *factorización* de una matriz A es una ecuación que expresa a A como un producto de dos o más matrices. Mientras que la multiplicación de matrices implica una síntesis de datos (combinando el efecto de dos o más transformaciones lineales en una sola matriz), la factorización de matrices es un *análisis* de datos. En el lenguaje de la ciencia de las computadoras, la expresión de A como un producto equivale a un *preprocesamiento* de los datos de A, el cual organiza esos datos en dos o más partes cuyas estructuras son más útiles de algún modo, quizá por ser más accesibles para realizar cálculos con ellas.

Las factorizaciones de matrices y, después, las factorizaciones de transformaciones lineales aparecerán en un buen número de puntos clave a lo largo de este texto. Esta sección se enfoca en una factorización que es el centro de varios programas de computadora importantes usados de manera extensa en aplicaciones. Algunas otras factorizaciones, que se estudiarán después, se presentan en los ejercicios.

La factorización LU

La factorización LU, descrita a continuación, está motivada por el muy frecuente problema industrial y de negocios que consiste en resolver una sucesión de ecuaciones, todas con la misma matriz de coeficientes:

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad \dots, \quad A\mathbf{x} = \mathbf{b}_p \tag{1}$$

Vea el ejercicio 32, por ejemplo. También vea la sección 5.8, donde se usa el método de la potencia inversa para estimar los valores propios de una matriz resolviendo ecuaciones como (1), una a la vez.

Cuando A es invertible, se podría calcular A^{-1} y luego calcular $A^{-1}\mathbf{b}_1$, $A^{-1}\mathbf{b}_2$, y así sucesivamente. Sin embargo, resulta más eficiente resolver la primera ecuación de (1) mediante reducción por filas y obtener una factorización LU de A al mismo tiempo. Después, las ecuaciones restantes de (1) se resuelven con la factorización LU.

Suponga de inicio que A es una matriz $m \times n$ que puede reducirse a su forma escalonada *sin intercambios de fila.* (Después se estudiará e1 caso general.) Entonces A puede escribirse en la forma $A = LU$, donde L es una matriz triangular inferior de $m \times m$ con números 1 en la diagonal y U es una forma escalonada de $m \times n$ de A. Por ejemplo, vea la figura 1. Una factorización de este tipo se llama **factorización LU** de A. La matriz L es invertible y se denomina matriz triangular inferior *unitaria*.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix} \begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\qquad L \qquad\qquad\qquad\qquad U$$

FIGURA 1 Una factorización LU.

Antes de estudiar la forma de construir L y U, es necesario examinar la razón de su utilidad. Cuando $A = LU$, la ecuación $A\mathbf{x} = \mathbf{b}$ se puede escribir como $L(U\mathbf{x}) = \mathbf{b}$. Escribiendo $\mathbf{y}$ en lugar de $U\mathbf{x}$, se puede encontrar $\mathbf{x}$ al resolver el *par* de ecuaciones

$$\boxed{\begin{aligned} L\mathbf{y} &= \mathbf{b} \\ U\mathbf{x} &= \mathbf{y} \end{aligned}} \tag{2}$$

Primero se despeja $\mathbf{y}$ de $L\mathbf{y} = \mathbf{b}$, y luego se resuelve $U\mathbf{x} = \mathbf{y}$ para obtener $\mathbf{x}$. Vea la figura 2. Las dos ecuaciones resultan fáciles de resolver porque L y U son triangulares.

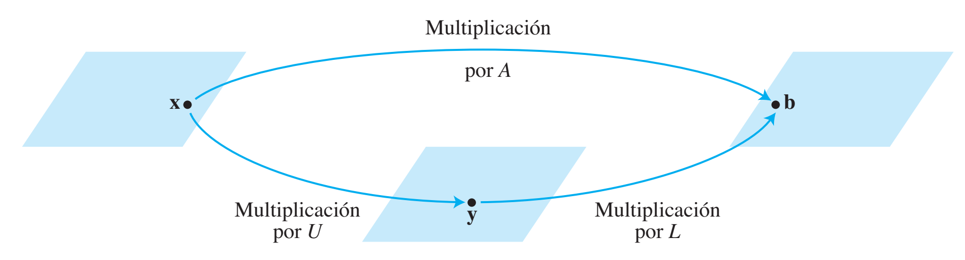


FIGURA 2 Factorización de la función $\mathbf{x} \mapsto A\mathbf{x}$.

EJEMPLO 1 Se puede verificar que

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

Use esta factorización LU para resolver $A\mathbf{x} = \mathbf{b}$, donde $\mathbf{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$.

Solución La resolución de $L\mathbf{y} = \mathbf{b}$ requiere únicamente de 6 multiplicaciones y 6 sumas, porque la aritmética ocurre sólo en la columna 5. (En L, los ceros debajo de cada pivote se crean automáticamente con la elección de las operaciones por fila.)

$$\begin{bmatrix} L & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} I & \mathbf{y} \end{bmatrix}$$

Entonces, para $U\mathbf{x} = \mathbf{y}$, la etapa "regresiva" de la reducción por filas requiere de 4 divisiones, 6 multiplicaciones y 6 sumas. (Por ejemplo, para producir la columna 4 de $[U \quad \mathbf{y}]$ se requieren una división en la fila 4 y tres pares multiplicación-suma para sumar múltiplos de la fila 4 a las filas de arriba.)

$$\begin{bmatrix} U & \mathbf{y} \end{bmatrix} = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

Para encontrar $\mathbf{x}$ se requieren 28 operaciones aritméticas, u operaciones de punto flotante ("flops"), sin contar el costo de encontrar L y U. En contraste, la reducción por filas de $[A \quad \mathbf{b}]$ hasta $[I \quad \mathbf{x}]$ requiere de 62 operaciones. ■

La eficiencia computacional de la factorización LU depende de que se conozcan L y U. El siguiente algoritmo muestra que la reducción por filas de A a su forma escalonada U equivale a una factorización LU, porque produce L prácticamente sin trabajo extra. Después de la primera reducción por filas, L y U se obtienen al resolver ecuaciones adicionales cuya matriz de coeficientes es A.

Un algoritmo de factorización LU

Suponga que A puede reducirse a una forma escalonada U empleando sólo reemplazos de filas que suman un múltiplo de una fila a otra situada *debajo* de la primera. En este caso, existen matrices elementales triangulares inferiores unitarias $E_1, \ldots, E_p$ tales que

$$E_p \cdots E_1 A = U \tag{3}$$

Entonces

$$A = (E_p \cdots E_1)^{-1} U = LU$$

donde

$$L = (E_p \cdots E_1)^{-1} \tag{4}$$

Puede demostrarse que los productos y los inversos de las matrices triangulares inferiores unitarias son también triangulares inferiores unitarios. (Por ejemplo, vea el ejercicio 19.) Así, L es triangular inferior unitaria.

Observe que las operaciones por fila en (3), que reducen A a U, también reducen la L de (4) a I, debido a que $E_p \cdots E_1 L = (E_p \cdots E_1)(E_p \cdots E_1)^{-1} = I$. Esta observación es la clave para *construir* L.

ALGORITMO PARA UNA FACTORIZACIÓN LU

1. Reduzca A a una forma escalonada U mediante una sucesión de operaciones de reemplazo de filas, si esto es posible.
2. Coloque las entradas de L de tal manera que la misma sucesión de operaciones por fila reduzca L a I.

El paso 1 no siempre es posible, pero cuando lo es, el argumento anterior muestra que existe una factorización LU. En el ejemplo 2 se mostrará cómo implementar el paso 2. Por construcción, L satisfará

$$(E_p \cdots E_1)L = I$$

donde se usan las mismas $E_1, \dots, E_p$ que en (3). Así, L será invertible, por el teorema de la matriz invertible, con $(E_p \cdots E_1) = L^{-1}$. A partir de (3), $L^{-1}A = U$, y $A = LU$. Por lo tanto, el paso 2 producirá una L aceptable.

EJEMPLO 2 Encuentre una factorización LU de

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$$

Solución Dado que A tiene cuatro filas, L debe ser de 4×4. La primera columna de L es la primera columna de A dividida entre la entrada pivote superior:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & & 1 & 0 \\ -3 & & & 1 \end{bmatrix}$$

Compare las primeras columnas de A y de L. *Las operaciones por fila que crearon ceros en la primera columna de A también crearán ceros en la primera columna de L.* Se desea que esta misma correspondencia de operaciones por fila sea válida para el resto de L, así que se examina una reducción por filas de A a una forma escalonada U:

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} = A_1 \quad (5)$$

$$\sim A_2 = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U$$

Las entradas resaltadas determinan la reducción por filas de A a U. En cada pivote, divida las entradas resaltadas entre el pivote y coloque el resultado en L:

$$\begin{bmatrix} 2 \\ -4 \\ 2 \\ -6 \end{bmatrix} \begin{bmatrix} 3 \\ -9 \\ 12 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 5 \end{bmatrix}$$

$$\div 2 \quad \div 3 \quad \div 2 \quad \div 5$$

$$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

$$\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix}, \quad \text{y} \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}$$

Con un cálculo sencillo puede verificarse que estas L y U satisfacen $LU = A$. ■

En el trabajo práctico, casi siempre son necesarios los intercambios de fila, porque se usa el pivoteo parcial para lograr una precisión alta. (Recuerde que este procedimiento selecciona, entre las posibles opciones de pivote, una entrada en la columna que tenga el mayor valor absoluto.) Para manejar los intercambios de fila, la factorización LU anterior puede modificarse con facilidad para producir una L que es *triangular inferior permutada*, en el sentido de que un arreglo (llamado permutación) de las filas de L puede hacer que L sea triangular inferior (unitaria). La *factorización LU permutada* resultante resuelve $A\mathbf{x} = \mathbf{b}$ en la misma forma que antes, excepto que la reducción de $[L \quad \mathbf{b}]$ a $[I \quad \mathbf{y}]$ sigue el orden de los pivotes de L de izquierda a derecha, empezando con el pivote de la primera columna. Una referencia a una “factorización LU” incluye, por lo general, la posibilidad de que L pueda ser triangular inferior permutada. Para mayores detalles, vea la *guía de estudio* (*Study Guide*).

SG Factorizaciones LU permutadas 2 a 24 (Permuted LU Factorizations 2-24)

NOTAS NUMÉRICAS

Los siguientes conteos de operaciones corresponden a una matriz densa A de $n \times n$ (con la mayor parte de sus entradas distintas de cero), donde n es moderadamente grande, por ejemplo, $n \geq 30$.[1]

1. El cálculo de una factorización LU de A requiere $2n^3/3$ flops (aproximadamente lo mismo que reducir por filas $[A \quad \mathbf{b}]$), mientras que encontrar A^{-1} demanda alrededor de $2n^3$ flops.
2. La resolución de $L\mathbf{y} = \mathbf{b}$ y $U\mathbf{x} = \mathbf{y}$ requiere alrededor de $2n^2$ flops, debido a que cualquier sistema triangular $n \times n$ puede resolverse en aproximadamente n^2 flops.
3. La multiplicación de $\mathbf{b}$ por A^{-1} también requiere cerca de $2n^2$ operaciones, pero podría ser que el resultado no sea tan preciso como el obtenido a partir de L y U (debido al error de redondeo cuando se calculan tanto A^{-1} como $A^{-1}\mathbf{b}$).
4. Si A es dispersa (la mayor parte de sus entradas son cero), entonces L y U podrían ser dispersas también, pero es probable que A^{-1} sea densa. En este caso, resulta *mucho* más rápido resolver $A\mathbf{x} = \mathbf{b}$ con una factorización LU que usar A^{-1}. Vea el ejercicio 31.

CD Operaciones de punto flotante (Floating Point Operations)

[1]Vea la sección 3.8 de *Applied Linear Algebra*, 3a. ed., de Ben Noble y James W. Daniel (Englewood Cliffs, NJ: Prentice-Hall, 1988). Recuerde que para los propósitos de este curso, un *flop* es $+$, $-$, $\times$, o $\div$.

Una factorización de matrices en ingeniería eléctrica

La factorización de matrices está íntimamente relacionada con el problema de construir una red eléctrica de propiedades específicas. El análisis que se presenta enseguida permite vislumbrar la conexión entre factorización y diseño de circuitos.

Suponga que la caja de la figura 3 representa algún tipo de circuito eléctrico, con una entrada y una salida. El voltaje y la corriente de entrada se registran como $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$ (con el voltaje v en volts y la corriente i en amperes), y el voltaje y la corriente de salida se registran como $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$. Es frecuente que la transformación $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} \mapsto \begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$ sea lineal. Esto es, existe una matriz A, llamada *matriz de transferencia*, tal que

$$\begin{bmatrix} v_2 \\ i_2 \end{bmatrix} = A \begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$$

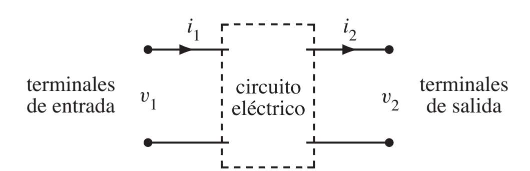


FIGURA 3 Un circuito con terminales de entrada y salida.

En la figura 4 se muestra una *red en escalera*, donde dos circuitos (podría haber más) están conectados en serie, de modo que la salida de un circuito sea la entrada del siguiente circuito. El circuito de la izquierda en la figura 4 es un *circuito en serie*, con resistencia R_1 (en ohms).

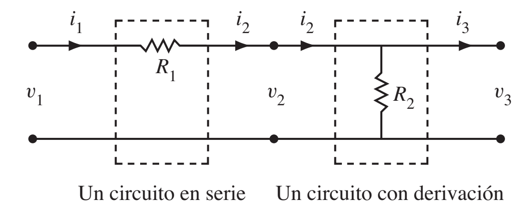


FIGURA 4 Una red en escalera.

El circuito de la derecha en la figura 4 es un *circuito con derivación*, y resistencia R_2. Si se utiliza la ley de Ohm y las leyes de Kirchhoff, puede demostrarse que las matrices de transferencia de los circuitos en serie y con derivación son, respectivamente,

$$\underset{\text{Matriz de transferencia del circuito en serie}}{\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix}} \quad \text{y} \quad \underset{\text{Matriz de transferencia del circuito con derivación}}{\begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}}$$

EJEMPLO 3

a. Encuentre la matriz de transferencia para la red en escalera de la figura 4.

b. Diseñe una red en escalera cuya matriz sea $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$.

Solución

a. Sean A_1 y A_2 las matrices de transferencia de los circuitos en serie y con derivación, respectivamente. Entonces, un vector de entrada $\mathbf{x}$ se transforma primero en $A_1\mathbf{x}$ y luego en $A_2(A_1\mathbf{x})$. La conexión en serie de los circuitos corresponde a la composición de transformaciones lineales, y la matriz de transferencia de la red en escalera es (observe el orden)

$$A_2A_1 = \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} \quad (6)$$

b. Se pretende factorizar la matriz $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$ para obtener el producto de matrices de transferencia, como en (6). Así que se buscan las R_1 y R_2 de la figura 4 que satisfagan

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$$

De las entradas (1, 2), se tiene que $R_1 = 8$ ohms, y de las entradas (2, 1), $1/R_2 = .5$ ohms y $R_2 = 1/.5 = 2$ ohms. Con estos valores, la red de la figura 4 tiene la matriz de transferencia deseada.

Una matriz de transferencia de red resume el comportamiento de entrada-salida (las especificaciones de diseño) de la red, sin referencia a los circuitos internos. Para construir físicamente una red con propiedades específicas, un ingeniero determina al principio si dicha red puede construirse (o *realizarse*). Después, trata de factorizar la matriz de transferencia para obtener matrices correspondientes a circuitos más pequeños que quizá ya fueron fabricados y estén listos para ensamblarse. En el caso frecuente de la corriente alterna, las entradas de la matriz de transferencia normalmente son funciones con valores complejos. (Vea los ejercicios 19 y 20 de la sección 2.4 y el ejemplo 2 de la sección 3.3.) Un problema estándar consiste en encontrar una *realización mínima* que use el menor número de componentes eléctricos.

PROBLEMA DE PRÁCTICA

Encuentre una factorización LU de $A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$. [*Nota:* Resultará que A tiene solamente tres columnas pivote, entonces el método del ejemplo 2 sólo produce las tres primeras columnas de L. Las dos columnas restantes de L provienen de I_5.]

2.5 EJERCICIOS

En los ejercicios 1 a 6, resuelva la ecuación $A\mathbf{x} = \mathbf{b}$ usando la factorización LU dada para A. En los ejercicios 1 y 2, resuelva también $A\mathbf{x} = \mathbf{b}$ por reducción ordinaria de columnas.

1. $A = \begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -7 \\ 5 \\ 2 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -5 & 1 \end{bmatrix}\begin{bmatrix} 3 & -7 & -2 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \end{bmatrix}$$

2. $A = \begin{bmatrix} 4 & 3 & -5 \\ -4 & -5 & 7 \\ 8 & 6 & -8 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & 3 & -5 \\ 0 & -2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$

3. $A = \begin{bmatrix} 2 & -1 & 2 \\ -6 & 0 & -2 \\ 8 & -1 & 5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & -1 & 2 \\ 0 & -3 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

4. $A = \begin{bmatrix} 2 & -2 & 4 \\ 1 & -3 & 1 \\ 3 & 7 & 5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -5 & 1 \end{bmatrix}\begin{bmatrix} 2 & -2 & 4 \\ 0 & -2 & -1 \\ 0 & 0 & -6 \end{bmatrix}$$

5. $A = \begin{bmatrix} 1 & -2 & -4 & -3 \\ 2 & -7 & -7 & -6 \\ -1 & 2 & 6 & 4 \\ -4 & -1 & 9 & 8 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 7 \\ 0 \\ 3 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -4 & 3 & -5 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 & -4 & -3 \\ 0 & -3 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

6. $A = \begin{bmatrix} 1 & 3 & 4 & 0 \\ -3 & -6 & -7 & 2 \\ 3 & 3 & 0 & -4 \\ -5 & -3 & 2 & 9 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ -5 & 4 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 & 4 & 0 \\ 0 & 3 & 5 & 2 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Encuentre una factorización LU de las matrices de los ejercicios 7 a 6 (con L triangular inferior unitaria). Observe que MATLAB generalmente produce una factorización LU permutada porque utiliza pivoteo parcial para lograr exactitud numérica.

7. $\begin{bmatrix} 2 & 5 \\ -3 & -4 \end{bmatrix}$

8. $\begin{bmatrix} 6 & 9 \\ 4 & 5 \end{bmatrix}$

9. $\begin{bmatrix} 3 & -1 & 2 \\ -3 & -2 & 10 \\ 9 & -5 & 6 \end{bmatrix}$

10. $\begin{bmatrix} -5 & 3 & 4 \\ 10 & -8 & -9 \\ 15 & 1 & 2 \end{bmatrix}$

11. $\begin{bmatrix} 3 & -6 & 3 \\ 6 & -7 & 2 \\ -1 & 7 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 2 & -4 & 2 \\ 1 & 5 & -4 \\ -6 & -2 & 4 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 3 & -5 & -3 \\ -1 & -5 & 8 & 4 \\ 4 & 2 & -5 & -7 \\ -2 & -4 & 7 & 5 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 4 & -1 & 5 \\ 3 & 7 & -2 & 9 \\ -2 & -3 & 1 & -4 \\ -1 & 6 & -1 & 7 \end{bmatrix}$

15. $\begin{bmatrix} 2 & -4 & 4 & -2 \\ 6 & -9 & 7 & -3 \\ -1 & -4 & 8 & 0 \end{bmatrix}$

16. $\begin{bmatrix} 2 & -6 & 6 \\ -4 & 5 & -7 \\ 3 & 5 & -1 \\ -6 & 4 & -8 \\ 8 & -3 & 9 \end{bmatrix}$

17. Cuando A es invertible, MATLAB encuentra A^{-1} al factorizar $A = LU$ (donde L puede ser triangular inferior permutada), invirtiendo L y U, y calculando luego $U^{-1}L^{-1}$. Use este método para calcular el inverso de A en el ejercicio 2. (Aplique el algoritmo de la sección 2.2 a L y a U.)

18. Encuentre A^{-1} como en el ejercicio 17, usando A del ejercicio 3.

19. Sea A una matriz $n \times n$ triangular inferior con entradas diferentes de cero en la diagonal. Demuestre que A es invertible y A^{-1} es triangular inferior. [*Indicación:* Explique por qué A puede convertirse en I usando sólo reemplazos de filas y escalamientos. (¿Dónde están los pivotes?) También, explique por qué las operaciones por filas que reducen A a I transforman I en una matriz triangular inferior.]

20. Sea $A = LU$ una factorización LU. Explique por qué A puede reducirse por filas a U empleando solamente operaciones de reemplazo. (Este hecho es el recíproco de lo que se demostró en el texto.)

21. Suponga $A = BC$, donde B es invertible. Muestre que cualquier sucesión de operaciones por fila que reduzca B a I tam-

bién reduce A a C. El recíproco no es cierto, puesto que la matriz cero puede factorizarse como $0 = B \cdot 0$.

Los ejercicios 22 a 26 proporcionan una visualización de ciertas factorizaciones de matriz ampliamente utilizadas; algunas de las cuales se discuten posteriormente en el texto.

22. (*Factorización LU reducida.*) Con A como en el problema de práctica, encuentre una matriz B de 5×3 y una matriz C de 3×4 tales que $A = BC$. Generalice esta idea para caso donde A es $m \times n$, $A = LU$, y U tiene solamente tres filas diferentes de cero.

23. (*Factorización de rango.*) Suponga que una matriz A admite una factorización $A = CD$, donde C es de $m \times 4$ y D de $4 \times n$.

a. Demuestre que A es la suma de cuatro productos exteriores. (Vea la sección 2.4.)

b. Sea $m = 400$ y $n = 100$. Explique por qué un programador de computadoras podría preferir almacenar los datos de A en forma de dos matrices C y D.

24. (*Factorización QR.*) Suponga que $A = QR$, donde Q y R son $n \times n$, R es invertible y triangular superior, y Q tiene la propiedad de que $Q^TQ = I$. Demuestre que para cada $\mathbf{b}$ en $\mathbb{R}^n$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución única. ¿Cuáles cálculos con Q y R producirían la solución?

25. (*Descomposición en valores singulares.*) Suponga que $A = UDV^T$, donde U y V son matrices de $n \times n$ con la propiedad de que $U^TU = I$ y $V^TV = I$, y donde D es una matriz diagonal con números positivos $\sigma_1, \ldots, \sigma_n$ sobre la diagonal. Muestre que A es invertible y encuentre una fórmula para A^{-1}.

26. (*Factorización espectral.*) Suponga que una matriz A de 3×3 admite una factorización como $A = PDP^{-1}$, donde P es alguna matriz invertible de 3×3 y D es la matriz diagonal

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$$

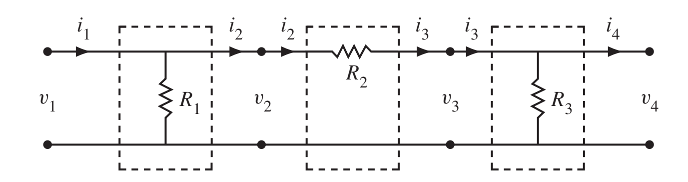
Muestre que esta factorización es útil cuando se calculan potencias altas de A. Encuentre fórmulas relativamente sencillas para A^2, A^3 y A^k (k es un entero positivo), usando P y las entradas en D.

27. Diseñe dos redes en escalera diferentes con salida de 9 volts y 4 amperes cuando la entrada sea de 12 volts y 6 amperes.

28. Muestre que si tres circuitos de derivación (con resistencias R_1, R_2, R_3) se conectan en serie, la red resultante tiene la misma matriz de transferencia que un único circuito con derivación. Encuentre una fórmula para la resistencia que haya en ese circuito.

29. a. Encuentre la matriz de transferencia de la red que se muestra en la figura.

b. Sea $A = \begin{bmatrix} 4/3 & -12 \\ -1/4 & 3 \end{bmatrix}$. Diseñe una red en escalera cuya matriz de transferencia sea A encontrando una factorización matricial apropiada de A.

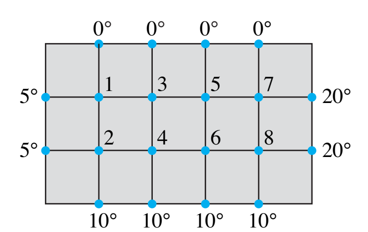


30. Encuentre una factorización diferente de la A del ejercicio 29, y a partir de ella diseñe una red en escalera distinta cuya matriz de transferencia sea A.

31. **[M]** La solución al problema de flujo de calor en estado estable para la placa de la figura se aproxima al resolver la ecuación $A\mathbf{x} = \mathbf{b}$, donde $\mathbf{b} = (5, 15, 0, 10, 0, 10, 20, 30)$ y

$$A = \begin{bmatrix} 4 & -1 & -1 & & & & & \\ -1 & 4 & 0 & -1 & & & & \\ -1 & 0 & 4 & -1 & -1 & & & \\ & -1 & -1 & 4 & 0 & -1 & & \\ & & -1 & 0 & 4 & -1 & -1 & \\ & & & -1 & -1 & 4 & 0 & -1 \\ & & & & -1 & 0 & 4 & -1 \\ & & & & & -1 & -1 & 4 \end{bmatrix}$$

WEB



(Remítase al ejercicio 35 de la sección 1.1.) Las entradas faltantes en A son ceros. Las entradas diferentes de cero de A quedan dentro de una banda a lo largo de la diagonal principal. Tales *matrices de banda* se dan en diversas aplicaciones, y a menudo son extremadamente grandes (con miles de filas y columnas pero con bandas relativamente angostas).

a. Encuentre la factorización LU de A, y observe que ambos factores son matrices de banda (con dos diagonales diferentes de cero abajo o arriba de la diagonal principal). Calcule $LU - A$ para comprobar su trabajo.

b. Use la factorización LU para resolver $A\mathbf{x} = \mathbf{b}$.

c. Obtenga A^{-1} y observe que A^{-1} es una matriz densa sin estructura de banda. Cuando A es grande, L y U pueden almacenarse en mucho menos espacio que A^{-1}. Este hecho es otra razón para preferir la factorización LU de A en lugar de la propia A^{-1}.

32. **[M]** La matriz de banda A que se muestra a continuación puede servir para calcular la conducción de calor no estacionaria en una barra para la cual las temperaturas en sus puntos $p_1, \ldots, p_5$ cambian con el tiempo.[2]

Δx Δx

p_1 p_2 p_3 p_4 p_5

La constante C de la matriz depende de la naturaleza física de la barra, de la distancia Δx entre los puntos de la barra, y del tiempo Δt que transcurra entre mediciones sucesivas de temperatura. Suponga que para $k = 0, 1, 2, \ldots$, un vector $\mathbf{t}_k$ en $\mathbb{R}^5$ enlista las temperaturas en el tiempo $k\Delta t$. Si ambos extremos de la barra se mantienen a 0°, entonces los vectores de temperatura satisfacen la ecuación $A\mathbf{t}_{k+1} = \mathbf{t}_k$ ($k = 0, 1, \ldots$), donde

$$A = \begin{bmatrix} (1+2C) & -C & & & \\ -C & (1+2C) & -C & & \\ & -C & (1+2C) & -C & \\ & & -C & (1+2C) & -C \\ & & & -C & (1+2C) \end{bmatrix}$$

a. Encuentre la factorización LU de A cuando $C = 1$. Una matriz como A con tres diagonales diferentes de cero se denomina *matriz tridiagonal*. Los factores L y U son *matrices bidiagonales*.

b. Suponga que $C = 1$ y $\mathbf{t}_0 = (10, 12, 12, 12, 10)$. Use la factorización LU de A para encontrar las distribuciones de temperatura $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$ y $\mathbf{t}_4$.

[2] Vea Biswa N. Datta, *Numerical Linear Algebra and Applications* (Pacific Grove, CA: Brooks/Cole, 1994), págs. 200-201.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

Divida las entradas de cada columna resaltada mediante el pivote en la parte superior. Las columnas resultantes forman las tres primeras columnas de la mitad inferior de L. Esto basta para hacer que la reducción por filas de L a I corresponda a la reducción de A a U. Use las dos últimas columnas de I_5 para hacer a L triangular inferior unitaria.

$$\begin{bmatrix} 2 \\ 6 \\ 2 \\ 4 \\ -6 \end{bmatrix} \quad \begin{bmatrix} 3 \\ -3 \\ 6 \\ -9 \end{bmatrix} \quad \begin{bmatrix} 5 \\ -5 \\ 10 \end{bmatrix}$$

$$\div 2 \qquad \div 3 \qquad \div 5$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$\begin{bmatrix} 1 & & & \\ 3 & 1 & & \\ 1 & -1 & 1 & \cdots \\ 2 & 2 & -1 & \\ -3 & -3 & 2 & \end{bmatrix}, \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 2 & 2 & -1 & 1 & 0 \\ -3 & -3 & 2 & 0 & 1 \end{bmatrix}$$

2.6 EL MODELO DE LEONTIEF DE ENTRADA Y SALIDA

El álgebra lineal desempeñó un papel fundamental en el trabajo ganador del Premio Nobel de Wassily Leontief, como se mencionó al principio del capítulo 1. El modelo económico descrito en esta sección es la base de modelos más complejos usados actualmente en muchas partes del mundo.

Suponga que la economía de una nación se divide en n sectores que producen bienes o servicios, y sea $\mathbf{x}$ un **vector de producción** en $\mathbb{R}^n$ que enlista lo producido por cada sector en un año. También, suponga que otra parte de la economía (llamada *sector abierto*) no produce bienes ni servicios sino que solamente los consume, y sea $\mathbf{d}$ un **vector de demanda final** (o **relación de demandas finales**) que enlista los valores de los bienes y servicios demandados a los diversos sectores por la parte no productiva de la economía. El vector $\mathbf{d}$ puede representar la demanda del consumidor, el consumo del gobierno, la producción sobrante, las exportaciones, u otras demandas externas.

Conforme los diversos sectores producen bienes para satisfacer la demanda del consumidor, los productores crean por sí mismos una **demanda intermedia** adicional de bienes que necesitan como insumos para su propia producción. Las interrelaciones de los sectores son muy complejas, y la conexión entre la demanda final y la producción no es clara. Leontief se preguntó si hay un nivel de producción $\mathbf{x}$ tal que las cantidades producidas (o "suministradas") equilibren exactamente la demanda total de esa producción, de modo que

$$\left\{\begin{matrix}\text{cantidad}\\ \text{producida}\\ \mathbf{x}\end{matrix}\right\} = \left\{\begin{matrix}\text{demanda}\\ \text{intermedia}\end{matrix}\right\} + \left\{\begin{matrix}\text{demanda}\\ \text{final}\\ \mathbf{d}\end{matrix}\right\} \tag{1}$$

El supuesto básico del modelo de Leontief de entrada y salida es que, para cada sector, hay un **vector de consumo unitario** en $\mathbb{R}^n$ que enlista los insumos necesarios *por unidad de producción* del sector. Todas las unidades de entrada y salida se miden en millones de dólares, en lugar de cantidades como toneladas o fanegas. (Los precios de los bienes y servicios se mantienen constantes.)

Como un ejemplo simple, suponga que la economía consiste en tres sectores —manufactura, agricultura y servicios— con los vectores unitarios de consumo $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$ mostrados en la tabla siguiente:

	Insumos consumidos por unidad de producción		
Comprado por:	Manufactura	Agricultura	Servicios
Manufactura	.50	.40	.20
Agricultura	.20	.30	.10
Servicios	.10	.10	.30
	↑ $\mathbf{c}_1$	↑ $\mathbf{c}_2$	↑ $\mathbf{c}_3$

EJEMPLO 1 ¿Qué cantidades consumirá el sector de manufactura si decide producir 100 unidades?

Solución Calcule

$$100\mathbf{c}_1 = 100\begin{bmatrix}.50\\ .20\\ .10\end{bmatrix} = \begin{bmatrix}50\\ 20\\ 10\end{bmatrix}$$

Para producir 100 unidades, manufactura ordenará (es decir, “demandará”) y consumirá 50 unidades de otras partes del sector de manufactura, 20 unidades de agricultura, y 10 unidades de servicios.

Si manufactura decide producir x_1 unidades, entonces $x_1\mathbf{c}_1$ representa las *demandas intermedias* de manufactura, porque las cantidades de $x_1\mathbf{c}_1$ se consumirán en el proceso de creación de las x_1 unidades de producción. De la misma forma, si x_2 y x_3 denotan las produciones planeadas de los sectores de agricultura y servicios, $x_2\mathbf{c}_2$ y $x_3\mathbf{c}_3$ enlistan las demandas intermedias correspondientes. La demanda intermedia total de los tres sectores está dada por

$$\begin{aligned}\{\text{demanda intermedia}\} &= x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + x_3\mathbf{c}_3 \\ &= C\mathbf{x} \qquad (2)\end{aligned}$$

donde C es la **matriz de consumo** $[\mathbf{c}_1 \quad \mathbf{c}_2 \quad \mathbf{c}_3]$, a saber,

$$C = \begin{bmatrix} .50 & .40 & .20 \\ .20 & .30 & .10 \\ .10 & .10 & .30 \end{bmatrix} \qquad (3)$$

Las ecuaciones (1) y (2) producen el modelo de Leontief.

EL MODELO DE LEONTIEF DE ENTRADA-SALIDA, O ECUACIÓN DE PRODUCCIÓN

$$\underset{\text{Cantidad producida}}{\mathbf{x}} = \underset{\text{Demanda intermedia}}{C\mathbf{x}} + \underset{\text{Demanda final}}{\mathbf{d}} \qquad (4)$$

Si se escribe $\mathbf{x}$ como $I\mathbf{x}$ y se utiliza álgebra de matrices, es posible reescribir (4):

$$\begin{aligned} I\mathbf{x} - C\mathbf{x} &= \mathbf{d} \\ (I - C)\mathbf{x} &= \mathbf{d} \qquad (5)\end{aligned}$$

EJEMPLO 2 Considere la economía cuya matriz de consumo está dada por (3). Suponga que la demanda final es de 50 unidades para manufactura, 30 unidades para agricultura, y 20 unidades para servicios. Encuentre el nivel de producción $\mathbf{x}$ que satisfará esta demanda.

Solución La matriz de coeficientes en (5) es

$$I - C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} .5 & .4 & .2 \\ .2 & .3 & .1 \\ .1 & .1 & .3 \end{bmatrix} = \begin{bmatrix} .5 & -.4 & -.2 \\ -.2 & .7 & -.1 \\ -.1 & -.1 & .7 \end{bmatrix}$$

Para resolver (5), reduzca por filas la matriz aumentada

$$\begin{bmatrix} .5 & -.4 & -.2 & 50 \\ -.2 & .7 & -.1 & 30 \\ -.1 & -.1 & .7 & 20 \end{bmatrix} \sim \begin{bmatrix} 5 & -4 & -2 & 500 \\ -2 & 7 & -1 & 300 \\ -1 & -1 & 7 & 200 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & 226 \\ 0 & 1 & 0 & 119 \\ 0 & 0 & 1 & 78 \end{bmatrix}$$

La última columna se redondea a la unidad más cercana. El área de manufactura debe producir aproximadamente 226 unidades, agricultura 119 unidades, y servicios únicamente 78 unidades.

Si la matriz $I - C$ es invertible, se puede aplicar el teorema 5 de la sección 2.2 al reemplazar A con $(I - C)$, y de la ecuación $(I - C)\mathbf{x} = \mathbf{d}$ obtener $\mathbf{x} = (I - C)^{-1}\mathbf{d}$. El teorema siguiente muestra que en la mayoría de los casos prácticos, $I - C$ *es* invertible y el vector de producción $\mathbf{x}$ es económicamente factible, en el sentido de que las entradas de $\mathbf{x}$ son no negativas.

En el teorema, el término **suma de columna** denota la suma de las entradas en una columna de una matriz. En circunstancias ordinarias, las sumas de columna de una matriz de consumo son menores que uno porque un sector debería requerir menos de una unidad de insumos para generar una unidad de producción.

TEOREMA 11

Sea C la matriz de consumo de una economía y sea $\mathbf{d}$ la demanda final. Si C y $\mathbf{d}$ tienen entradas no negativas, y si cada suma de columna de C es menor que uno, entonces $(I - C)^{-1}$ existe y el vector de producción

$$\mathbf{x} = (I - C)^{-1}\mathbf{d}$$

tiene entradas no negativas y es la solución única de

$$\mathbf{x} = C\mathbf{x} + \mathbf{d}$$

El análisis siguiente sugerirá por qué el teorema es cierto, y conducirá a una nueva manera de calcular $(I - C)^{-1}$.

Una fórmula para $(I - C)^{-1}$

Imagine que la demanda representada por $\mathbf{d}$ se propone a las distintas industrias al inicio del año, y que las industrias responden estableciendo sus niveles de producción en $\mathbf{x} = \mathbf{d}$, lo cual satisfará exactamente la demanda final. Conforme las industrias se preparan para producir $\mathbf{d}$, emiten órdenes solicitando materia prima y otros insumos. Esto crea una demanda intermedia de insumos de $C\mathbf{d}$.

Para satisfacer la demanda adicional de $C\mathbf{d}$, las industrias necesitarán como insumos adicionales las cantidades de $C(C\mathbf{d}) = C^2\mathbf{d}$. Por supuesto, esto crea una segunda ronda de demanda intermedia, y cuando las industrias deciden producir aún más para satisfacer esta nueva demanda, se crea una tercera ronda de demanda, a saber, $C(C^2\mathbf{d}) = C^3\mathbf{d}$, y así sucesivamente.

En teoría, es posible imaginar que este proceso continúa de manera indefinida, aunque en la vida real no ocurriría en una sucesión tan rígida de acontecimientos. Puede hacerse un diagrama de esta situación hipotética en la forma siguiente:

	Demanda que debe satisfacerse	Insumos necesarios para satisfacer esta demanda
Demanda final	$\mathbf{d}$	$C\mathbf{d}$
Demanda intermedia		
1a. ronda	$C\mathbf{d}$	$C(C\mathbf{d}) = C^2\mathbf{d}$
2a. ronda	$C^2\mathbf{d}$	$C(C^2\mathbf{d}) = C^3\mathbf{d}$
3a. ronda	$C^3\mathbf{d}$	$C(C^3\mathbf{d}) = C^4\mathbf{d}$
	$\vdots$	$\vdots$

El nivel de producción $\mathbf{x}$ que satisfará el total de esta demanda es

$$\begin{aligned}\mathbf{x} &= \mathbf{d} + C\mathbf{d} + C^2\mathbf{d} + C^3\mathbf{d} + \cdots \\ &= (I + C + C^2 + C^3 + \cdots)\mathbf{d} \qquad (6)\end{aligned}$$

Para entender (6), se utiliza la siguiente identidad algebraica:

$$(I - C)(I + C + C^2 + \cdots + C^m) = I - C^{m+1} \qquad (7)$$

Puede demostrarse que si las sumas de columna en C son todas menores que 1, entonces $I - C$ es invertible, C^m se aproxima a la matriz cero cuando m crece de manera arbitraria, e $I - C^{m+1} \to I$. (Este hecho es análogo al de que si un número positivo t es menor que 1, entonces $t^m \to 0$ conforme aumenta m.) Usando (7), se escribe

$$(I - C)^{-1} \approx I + C + C^2 + C^3 + \cdots + C^m \qquad (8)$$

cuando las sumas de columna de C son menores que 1.

Se interpreta que (8) significa que el miembro derecho puede acercarse a $(I - C)^{-1}$ tanto como se desee al hacer a m lo suficientemente grande.

En los modelos de entrada y salida reales, las potencias de la matriz de consumo se aproximan a la matriz cero con cierta rapidez. Así que (8) realmente proporciona una manera práctica de calcular $(I - C)^{-1}$. De igual forma, para cualquier $\mathbf{d}$, los vectores $C^m\mathbf{d}$ se aproximan al vector cero rápidamente, y (6) es una manera práctica de resolver $(I - C)\mathbf{x} = \mathbf{d}$. Si las entradas de C y $\mathbf{d}$ son no negativas, entonces (6) muestra que las entradas de $\mathbf{x}$ también son no negativas.

Importancia económica de las entradas de $(I - C)^{-1}$

Las entradas de $(I - C)^{-1}$ son significativas porque pueden servir para predecir cómo tendrá que cambiar la producción $\mathbf{x}$ conforme cambie la demanda final $\mathbf{d}$. De hecho, las entradas de la columna j de $(I - C)^{-1}$ son las cantidades *incrementadas* que los diversos sectores tendrán que producir para satisfacer *un incremento de 1 unidad* en la demanda final de producción del sector j. Vea el ejercicio 8.

NOTA NUMÉRICA

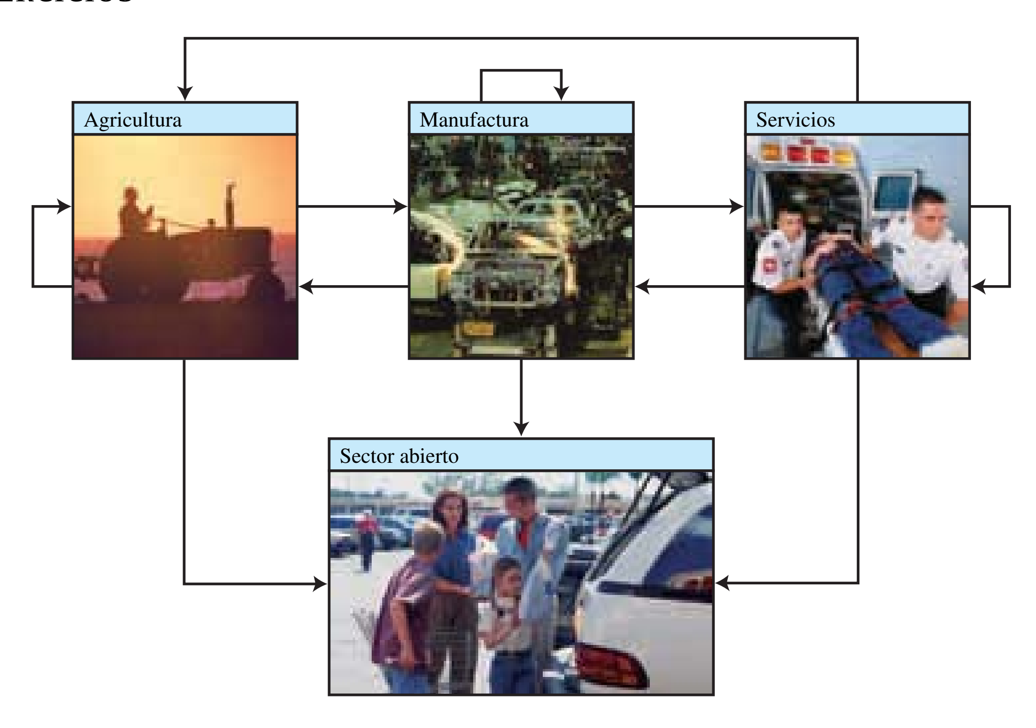
En cualquier problema de aplicación (no sólo en economía), una ecuación $A\mathbf{x} = \mathbf{b}$ puede escribirse siempre como $(I - C)\mathbf{x} = \mathbf{b}$, con $C = I - A$. Si el sistema es grande y *disperso* (con entradas cero en su mayoría), puede suceder que las sumas de columna de los valores absolutos en C sean menores que 1. En este caso, $C^m \to 0$. Si C^m tiende a cero con la suficiente rapidez, (6) y (8) proporcionarán fórmulas prácticas para resolver $A\mathbf{x} = \mathbf{b}$ y encontrar A^{-1}.

PROBLEMA DE PRÁCTICA

Suponga que una economía tiene dos sectores: bienes y servicios. Una unidad de producción de bienes requiere insumos de 0.2 unidades de bienes y 0.5 unidades de servicios. Una unidad de producción de servicios requiere insumos de 0.4 unidades de bienes y 0.3 unidades de servicios. Existe una demanda final de 20 unidades de bienes

y 30 unidades de servicios. Implemente el modelo de Leontief de entrada y salida para esta situación.

2.6 EJERCICIOS



Los ejercicios 1 a 4 se refieren a una economía dividida en tres sectores —manufactura, agricultura y servicios. Por cada unidad de producción, manufactura requiere de .10 unidades de otras compañías ubicadas en ese sector, de .30 unidades del sector agricultura, y de .30 unidades de servicios. Por cada unidad de producción, agricultura usa .20 unidades de su propia producción, .60 unidades de manufactura, y .10 unidades de servicios. Por cada unidad de producción el sector de servicios consume .10 unidades de servicios, .60 unidades de manufactura, pero ningún producto de agricultura.

1. Construya la matriz de consumo apropiada para esta economía, y determine cuáles demandas intermedias se crean si agricultura planea producir 100 unidades.
2. Determine los niveles de producción que se necesitan para satisfacer una demanda final de 18 unidades para agricultura, sin demanda final para los otros sectores. (No calcule una matriz inversa.)
3. Determine los niveles de producción necesarios para satisfacer una demanda final de 18 unidades para manufactura, sin demanda final para los otros sectores. (No calcule una matriz inversa.)
4. Determine los niveles de producción necesarios para satisfacer una demanda final de 18 unidades para manufactura, 18 para agricultura, y 0 unidades para servicios.
5. Considere el modelo de producción $\mathbf{x} = C\mathbf{x} + \mathbf{d}$ para una economía con dos sectores, donde

$$C = \begin{bmatrix} .0 & .5 \\ .6 & .2 \end{bmatrix}, \qquad \mathbf{d} = \begin{bmatrix} 50 \\ 30 \end{bmatrix}$$

   Use una matriz inversa y determine el nivel de producción necesario para satisfacer la demanda final.

6. Repita el ejercicio 5 con $C = \begin{bmatrix} .1 & .6 \\ .5 & .2 \end{bmatrix}$ y $\mathbf{d} = \begin{bmatrix} 18 \\ 11 \end{bmatrix}$.
7. Sean C y $\mathbf{d}$ como en el ejercicio 5.
   a. Determine el nivel de producción necesario para satisfacer una demanda final de una unidad de producción del sector 1.

b. Use una matriz inversa y determine el nivel de producción necesario para satisfacer una demanda final de $\begin{bmatrix} 51 \\ 30 \end{bmatrix}$.

c. Utilice el hecho de que $\begin{bmatrix} 51 \\ 30 \end{bmatrix} = \begin{bmatrix} 50 \\ 30 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ para explicar cómo y por qué están relacionadas entre sí las respuestas a los incisos (a), (b), y al ejercicio 5.

8. Sea C una matriz de consumo $n \times n$ cuyas sumas de columna son menores a 1. Sea $\mathbf{x}$ el vector de producción que satisface la demanda final $\mathbf{d}$, y sea $\Delta\mathbf{x}$ un vector de producción para satisfacer una demanda final diferente $\Delta\mathbf{d}$.

a. Muestre que si la demanda final cambia de $\mathbf{d}$ a $\mathbf{d} + \Delta\mathbf{d}$, entonces el nuevo nivel de producción debe ser $\mathbf{x} + \Delta\mathbf{x}$. Así, $\Delta\mathbf{x}$ proporciona las cantidades en que debe *cambiar* la producción para compensar el *cambio* $\Delta\mathbf{d}$ en la demanda.

b. Sea $\Delta\mathbf{d}$ el vector en $\mathbb{R}^n$ con 1 en la primera entrada y ceros en las demás entradas. Explique por qué la producción correspondiente $\Delta\mathbf{x}$ es la primera columna de $(I - C)^{-1}$. Esto muestra que la primera columna de $(I - C)^{-1}$ proporciona las cantidades que deben producir los diversos sectores para satisfacer un aumento de una unidad de la demanda final para la producción del sector 1.

9. Resuelva la ecuación de producción de Leontief para una economía con tres sectores, dado que

$$C = \begin{bmatrix} .2 & .2 & .0 \\ .3 & .1 & .3 \\ .1 & .0 & .2 \end{bmatrix} \quad \text{y} \quad \mathbf{d} = \begin{bmatrix} 40 \\ 60 \\ 80 \end{bmatrix}$$

10. La matriz de consumo C para la economía estadounidense en 1972 tiene la propiedad de que *cada entrada* de la matriz $(I - C)^{-1}$ es diferente de cero (y positiva).[1] ¿Qué dice esto acerca del efecto de aumentar la demanda de la producción en sólo un sector de la economía?

11. La ecuación de producción de Leontief, $\mathbf{x} = C\mathbf{x} + \mathbf{d}$, generalmente está acompañada por una **ecuación de precio** dual:

$$\mathbf{p} = C^T\mathbf{p} + \mathbf{v}$$

donde $\mathbf{p}$ es un **vector de precio** cuyas entradas enlistan el precio por unidad de producción de cada sector, y $\mathbf{v}$ es un **vector de valor agregado** cuyas entradas enlistan el valor agregado por unidad de producción. (El valor agregado incluye salarios, utilidades, depreciación, etc.) Un dato importante en economía es que el producto interno bruto (PIB) se puede expresar de dos maneras:

$$\{\text{producto interno bruto}\} = \mathbf{p}^T\mathbf{d} = \mathbf{v}^T\mathbf{x}$$

Verifique la segunda igualdad. [*Sugerencia:* Calcule $\mathbf{p}^T\mathbf{x}$ de dos maneras.]

12. Sea C una matriz de consumo tal que $C^m \to 0$ cuando $m \to \infty$, y para $m = 1, 2, \ldots$, sea $D_m = I + C + \cdots + C^m$. Encuentre una ecuación en diferencias que relacione D_m y D_{m+1} y, a partir de ella, obtenga un procedimiento iterativo para calcular la fórmula (8) para $(I - C)^{-1}$.

13. **[M]** La siguiente matriz de consumo C está basada en datos de entrada y salida para la economía de Estados Unidos en 1958, con datos para 81 sectores agrupados en 7 sectores más grandes: (1) productos no metálicos personales y domésticos, (2) productos metálicos finales (como vehículos de motor), (3) productos básicos de metal y minería, (4) productos básicos no metálicos y de agricultura, (5) energía, (6) servicios, y (7) entretenimiento y productos diversos.[2] Encuentre los niveles de producción necesarios para satisfacer la demanda final $\mathbf{d}$. (Las unidades están en millones de dólares.)

$$\begin{bmatrix} .1588 & .0064 & .0025 & .0304 & .0014 & .0083 & .1594 \\ .0057 & .2645 & .0436 & .0099 & .0083 & .0201 & .3413 \\ .0264 & .1506 & .3557 & .0139 & .0142 & .0070 & .0236 \\ .3299 & .0565 & .0495 & .3636 & .0204 & .0483 & .0649 \\ .0089 & .0081 & .0333 & .0295 & .3412 & .0237 & .0020 \\ .1190 & .0901 & .0996 & .1260 & .1722 & .2368 & .3369 \\ .0063 & .0126 & .0196 & .0098 & .0064 & .0132 & .0012 \end{bmatrix}$$

$$\mathbf{d} = \begin{bmatrix} 74{,}000 \\ 56{,}000 \\ 10{,}500 \\ 25{,}000 \\ 17{,}500 \\ 196{,}000 \\ 5{,}000 \end{bmatrix}$$

14. **[M]** El vector de demanda del ejercicio 13 es razonable para los datos de 1958, pero el análisis de Leontief de la economía mencionado en el mismo ejercicio utilizó un vector de demanda más cercano a los datos de 1964:

$$\mathbf{d} = (99640, 75548, 14444, 33501, 23527, 263985, 6526)$$

Encuentre los niveles de producción necesarios para satisfacer esta demanda.

15. **[M]** Use la ecuación (6) para resolver el problema del ejercicio 13. Establezca $\mathbf{x}^{(0)} = \mathbf{d}$, y para $k = 1, 2, \ldots$, calcule $\mathbf{x}^{(k)} = \mathbf{d} + C\mathbf{x}^{(k-1)}$. ¿Cuántos pasos se necesitan para obtener una respuesta al ejercicio 13 con cuatro cifras significativas?

---

[1]Wassily W. Leontief, "The World Economy of the Year 2000", *Scientific American*, septiembre de 1980, págs. 206-231.

[2]Wassily W. Leontief, "The Structure of the U.S. Economy", *Scientific American*, abril de 1965, págs. 30-32.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Se dan los siguientes datos:

Comprado de:	Insumos necesarios por unidad de producción: Bienes	Servicios	Demanda externa
Bienes	.2	.4	20
Servicios	.5	.3	30

El modelo de entrada y salida de Leontief es $\mathbf{x} = C\mathbf{x} + \mathbf{d}$, donde

$$C = \begin{bmatrix} .2 & .4 \\ .5 & .3 \end{bmatrix}, \qquad \mathbf{d} = \begin{bmatrix} 20 \\ 30 \end{bmatrix}$$

2.7 APLICACIONES A LOS GRÁFICOS POR COMPUTADORA

Los gráficos por computadora son imágenes desplegadas o animadas en una pantalla de computadora. Las aplicaciones de los gráficos por computadora están ampliamente difundidas y aumentan con rapidez. Por ejemplo, el diseño asistido por computadora (CAD, del inglés *computer-aided design*) es una parte integral de muchos procesos de ingeniería, tal como el proceso de diseño de aviones descrito en la introducción de este capítulo. La industria del entretenimiento ha realizado el uso más espectacular de los gráficos por computadora —desde los efectos especiales de *The Matrix* hasta la Xbox de PlayStation 2.

La mayor parte de los programas computacionales interactivos producidos para los negocios y la industria utiliza gráficos por computadora en los despliegues de pantalla y en otras funciones, como el despliegue gráfico de datos, la autoedición, y la producción de diapositivas para presentaciones comerciales y educativas. Por consiguiente, cualquier persona que estudie un lenguaje de computadora siempre pasa algún tiempo aprendiendo a usar gráficos de, por lo menos, dos dimensiones (2D).

En esta sección se examinará algo de las matemáticas básicas que se usan para manipular y desplegar imágenes gráficas tales como el modelo de alambre de un avión. Una imagen (o dibujo) de ese tipo consta de varios puntos, líneas rectas o curvas conectadas, e información sobre cómo llenar regiones cerradas delimitadas por esas líneas. A menudo, las líneas curvas se aproximan empleando segmentos de línea recta cortos, y una figura se define matemáticamente por medio de una lista de puntos.

Entre los símbolos gráficos más sencillos utilizados en 2D están las letras usadas como etiquetas en la pantalla. Algunas letras se guardan como objetos de alambre; otras que tienen porciones curvas se almacenan con fórmulas matemáticas adicionales para las curvas.

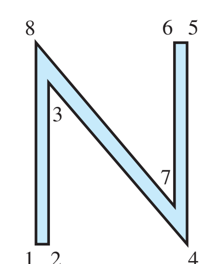


FIGURA 1
N regular.

EJEMPLO 1 La letra N mayúscula de la figura 1 está determinada por ocho puntos o *vértices.* Las coordenadas de los puntos pueden almacenarse en una matriz de datos, D.

$$\begin{array}{r} \\ \text{coordenada } x \\ \text{coordenada } y \end{array} \begin{array}{c} \text{Vértice:} \\ \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{bmatrix} 0 & .5 & .5 & 6 & 6 & 5.5 & 5.5 & 0 \\ 0 & 0 & 6.42 & 0 & 8 & 8 & 1.58 & 8 \end{bmatrix} \end{array} = D$$

Además de D, es necesario especificar cuáles vértices están conectados mediante líneas, pero aquí se omite este detalle.

La principal razón para describir los objetos gráficos por medio de segmentos de líneas rectas es que las transformaciones estándar en los gráficos de computadora mapean segmentos de línea sobre otros segmentos de línea. (Por ejemplo, vea el ejercicio 27 de la sección 1.8.) Una vez transformados los vértices que describen un objeto, se pueden conectar sus imágenes con las líneas rectas apropiadas para producir la imagen completa del objeto original.

EJEMPLO 2 Dada $A = \begin{bmatrix} 1 & .25 \\ 0 & 1 \end{bmatrix}$, describa el efecto de la transformación de trasquilado $\mathbf{x} \mapsto A\mathbf{x}$ sobre la letra N del ejemplo 1.

Solución Por la definición de multiplicación de matrices, las columnas del producto AD contienen las imágenes de los vértices de la letra N.

$$\begin{matrix} & \scriptstyle 1 & \scriptstyle 2 & \scriptstyle 3 & \scriptstyle 4 & \scriptstyle 5 & \scriptstyle 6 & \scriptstyle 7 & \scriptstyle 8 \end{matrix}$$
$$AD = \begin{bmatrix} 0 & .5 & 2.105 & 6 & 8 & 7.5 & 5.895 & 2 \\ 0 & 0 & 6.420 & 0 & 8 & 8 & 1.580 & 8 \end{bmatrix}$$

Los vértices transformados se grafican en la figura 2, junto con los segmentos de línea conectores que corresponden a los de la figura original.

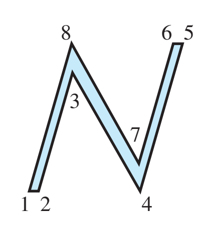
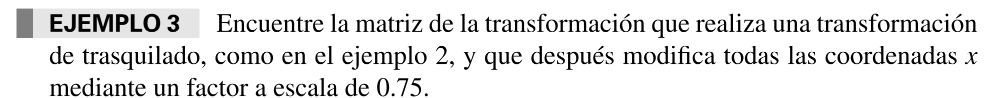


FIGURA 2
N inclinada.

La *N* cursiva de la figura 2 se ve demasiado ancha. Para compensarlo, se puede reducir la anchura mediante una transformación de escala.



EJEMPLO 3 Encuentre la matriz de la transformación que realiza una transformación de trasquilado, como en el ejemplo 2, y que después modifica todas las coordenadas x mediante un factor a escala de 0.75.

Solución La matriz que multiplica la coordenada x de un punto por 0.75 es

$$S = \begin{bmatrix} .75 & 0 \\ 0 & 1 \end{bmatrix}$$

Así que la matriz de la transformación compuesta es

$$\begin{aligned} SA &= \begin{bmatrix} .75 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & .25 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} .75 & .1875 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

El resultado de esta transformación compuesta se muestra en la figura 3.

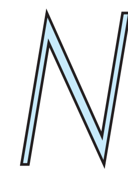
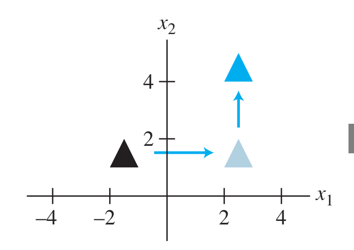


FIGURA 3
Transformación compuesta de *N*.

Las matemáticas de los gráficos por computadora están íntimamente conectadas con la multiplicación de matrices. Desafortunadamente, trasladar un objeto a una pantalla no corresponde directamente a la multiplicación de matrices porque la traslación no es una transformación lineal. La manera estándar de evitar esta dificultad es introducir lo que se conoce como *coordenadas homogéneas.*

Coordenadas homogéneas

Cada punto (x, y) en $\mathbb{R}^2$ puede identificarse con el punto $(x, y, 1)$ sobre el plano en $\mathbb{R}^3$ que se posiciona una unidad por encima del plano xy. Se dice que (x, y) tiene *coordenadas homogéneas* $(x, y, 1)$. Por ejemplo, el punto $(0, 0)$ tiene coordenadas homogéneas (0, 0,



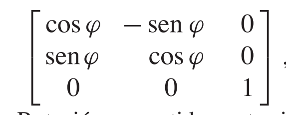
Traslación mediante $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$.

1). Las coordenadas homogéneas de puntos no se suman ni multiplican por escalares, pero se pueden transformar mediante multiplicación por matrices de 3×3.

EJEMPLO 4 Una traslación de la forma $(x, y) \mapsto (x + h, y + k)$ se escribe en coordenadas homogéneas como $(x, y, 1) \mapsto (x + h, y + k, 1)$. Esta transformación puede calcularse mediante multiplicación de matrices:

$$\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + h \\ y + k \\ 1 \end{bmatrix}$$

EJEMPLO 5 Cualquier transformación lineal sobre $\mathbb{R}^2$ se representa con respecto a coordenadas homogéneas por medio de una matriz partida de la forma $\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}$, donde A es una matriz de 2×2. Son ejemplos típicos:



$$\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} s & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Rotación en sentido contrario al de las manecillas del reloj con respecto al origen, ángulo φ — Reflexión a través de $y = x$ — Escalamiento de x mediante s y de y por medio de t

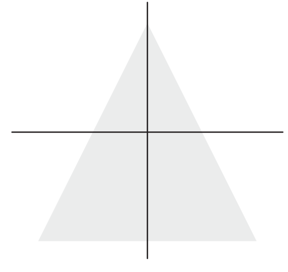
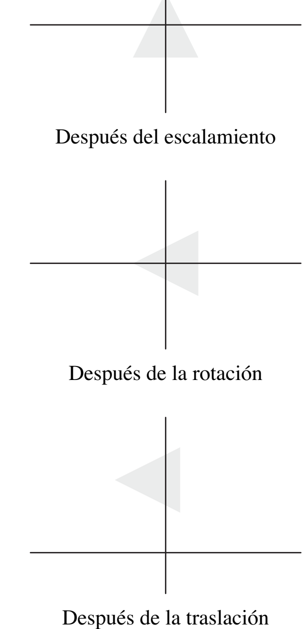


Figura original

Transformaciones compuestas

El movimiento de una figura en la pantalla de una computadora con frecuencia requiere utilizar dos o más transformaciones básicas. Cuando se usan coordenadas homogéneas, la composición de tales transformaciones corresponde a la multiplicación de matrices.



Después del escalamiento

Después de la rotación

Después de la traslación

EJEMPLO 6 Encuentre la matriz de 3×3 que corresponde a la transformación compuesta de aplicar un escalamiento por .3, una rotación de 90° y, por último, una traslación que suma $(-.5, 2)$ a cada punto de una figura.

Solución Si $\varphi = \pi/2$, entonces $\operatorname{sen}\varphi = 1$ y $\cos\varphi = 0$. A partir de los ejemplos se tiene que

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \xrightarrow{\text{Escalamiento}} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\xrightarrow{\text{Rotación}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\xrightarrow{\text{Traslación}} \begin{bmatrix} 1 & 0 & -.5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

La matriz para la transformación compuesta es

$$\begin{bmatrix} 1 & 0 & -.5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -1 & -.5 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -.3 & -.5 \\ .3 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

Gráficas tridimensionales por computadora

Algunos de los más recientes y estimulantes trabajos en gráficos por computadora se relacionan con el modelado molecular. Con gráficos tridimensionales (3D), un biólogo puede examinar una molécula simulada de proteína y buscar los sitios activos que pueden aceptar la molécula de un fármaco. El biólogo puede girar y trasladar un fármaco experimental para tratar de unirlo a la proteína. Esta capacidad de *visualizar* reacciones químicas potenciales es vital para la investigación de medicamentos modernos y del cáncer. De hecho, los avances en el diseño de medicamentos dependen, en cierta medida, del progreso que se logre en la capacidad de los gráficos por computadora para construir simulaciones realistas de las moléculas y sus interacciones.[1]

La investigación actual en el modelado de moléculas se enfoca en la *realidad virtual*, un entorno en el que un investigador puede ver y *sentir* la molécula de fármaco deslizarse dentro de la proteína. En la figura 4 se proporciona una retroalimentación táctil mediante un manipulador remoto que despliega la fuerza.

FIGURA 4 Modelado molecular en realidad virtual. (Departamento de Ciencias de la Computación, University of North Carolina en Chapel Hill. Fotografía de Bo Strain.)

[1]Robert Pool, "Computing in Science", *Science* **256**, 3 de abril de 1992, pág. 45.

Otro diseño de realidad virtual consiste en un casco y un guante que detectan movimientos de cabeza, mano y dedos. El casco incluye dos pequeñas pantallas de computadora, una para cada ojo. La búsqueda de que este entorno virtual sea más realista es un reto para ingenieros, científicos y matemáticos. Las matemáticas que se manejan aquí apenas abren la puerta a este campo de la investigación.

Coordenadas tridimensionales homogéneas

Por analogía con el caso bidimensional, se dice que $(x, y, z, 1)$ son las coordenadas homogéneas para el punto (x, y, z) en $\mathbb{R}^3$. En general, (X, Y, Z, H) son las **coordenadas homogéneas** para (x, y, z) si $H \neq 0$ y

$$x = \frac{X}{H}, \qquad y = \frac{Y}{H}, \qquad \text{y} \qquad z = \frac{Z}{H} \tag{1}$$

Cada múltiplo escalar diferente de cero de $(x, y, z, 1)$ proporciona un conjunto de coordenadas homogéneas para (x, y, z). Por ejemplo, $(10, -6, 14, 2)$ y $(-15, 9, -21, -3)$ son ambas coordenadas homogéneas para $(5, -3, 7)$.

El ejemplo siguiente ilustra las transformaciones que se usaron en el modelado molecular para introducir un fármaco en una molécula de proteína.

EJEMPLO 7 Construya matrices de 4×4 para las siguientes transformaciones:

a. Rotación con respecto al eje y en un ángulo de 30°. (Por convención, un ángulo positivo está en sentido contrario al de las manecillas del reloj cuando se ve hacia el origen desde la mitad positiva del eje de rotación, en este caso, el eje y.)

b. Traslación mediante el vector $\mathbf{p} = (-6, 4, 5)$.

Solución

a. Primero, construya la matriz de 3×3 para la rotación. El vector $\mathbf{e}_1$ gira hacia abajo en la dirección del eje z negativo, deteniéndose en $(\cos 30°, 0, -\text{sen } 30°) = (\sqrt{3}/2, 0, -.5)$. El vector $\mathbf{e}_2$ sobre el eje y no se mueve, pero $\mathbf{e}_3$ sobre el eje z gira hacia abajo en dirección del eje x positivo, deteniéndose en $(\text{sen } 30°, 0, \cos 30°) = (.5, 0, \sqrt{3}/2)$. Vea la figura 5. De la sección 1.9, la matriz estándar para esta rotación es

$$A = \begin{bmatrix} \sqrt{3}/2 & 0 & .5 \\ 0 & 1 & 0 \\ -.5 & 0 & \sqrt{3}/2 \end{bmatrix}$$

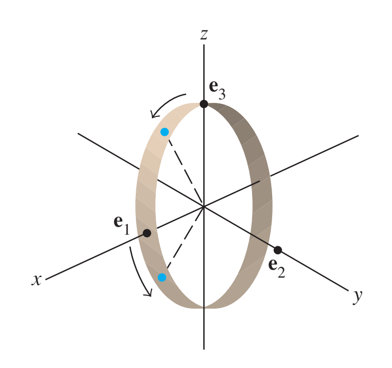


FIGURA 5

Por lo tanto, la matriz de rotación para las coordenadas homogéneas es

$$A = \begin{bmatrix} \sqrt{3}/2 & 0 & .5 & 0 \\ 0 & 1 & 0 & 0 \\ -.5 & 0 & \sqrt{3}/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

b. Se desea que $(x, y, z, 1)$ mapee a $(x - 6, y + 4, z + 5, 1)$. La matriz que hace esto es

$$\begin{bmatrix} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Proyecciones en perspectiva

Un objeto tridimensional se representa en la pantalla de dos dimensiones de una computadora proyectándolo sobre un *plano visual*. (Se omiten otros pasos importantes, como la selección de la porción del plano visual que se desplegará en la pantalla.) En aras de la simplicidad, considere que el plano xy representa la pantalla de la computadora, e imagine que el ojo de un observador está sobre el eje positivo z, en un punto $(0, 0, d)$. Una proyección *en perspectiva* mapea cada punto (x, y, z) sobre un punto de imagen $(x^*, y^*, 0)$ de manera que los dos puntos y la posición del ojo, llamada *centro de proyección*, estén sobre una línea. Vea la figura 6(a).

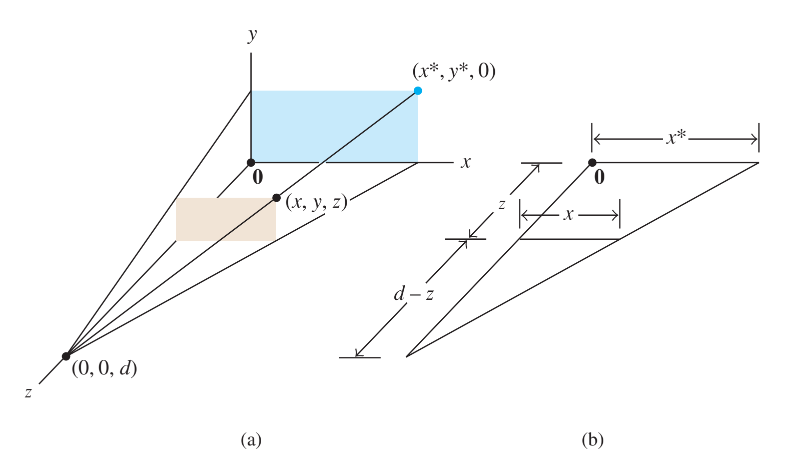


FIGURA 6 Proyección en perspectiva de (x, y, z) sobre $(x^*, y^*, 0)$.

El triángulo en el plano xz de la figura 6(a) se vuelve a trazar en la parte (b) mostrando la longitud de los segmentos de línea. Con triángulos similares se muestra que

$$\frac{x^*}{d} = \frac{x}{d - z} \quad \text{y} \quad x^* = \frac{dx}{d - z} = \frac{x}{1 - z/d}$$

De manera similar,

$$y^* = \frac{y}{1 - z/d}$$

Si se usan coordenadas homogéneas, es posible representar la proyección en perspectiva mediante una matriz, por ejemplo P. Se desea que $(x, y, z, 1)$ se transforme en $\left(\frac{x}{1-z/d}, \frac{y}{1-z/d}, 0, 1\right)$. Al escalar estas coordenadas por medio de $1 - z/d$, también puede utilizarse $(x, y, 0, 1 - z/d)$ como coordenadas homogéneas para la imagen. Ahora resulta fácil desplegar P. De hecho,

$$P\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/d & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \\ 1 - z/d \end{bmatrix}$$

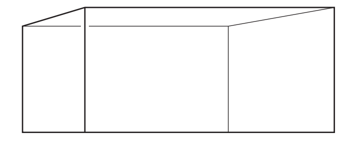
EJEMPLO 8 Sea S la caja con vértices (3, 1, 5), (5, 1, 5), (5, 0, 5), (3, 0, 5), (3, 1, 4), (5, 1, 4), (5, 0, 4) y (3, 0, 4). Encuentre la imagen de S bajo la proyección en perspectiva con centro de proyección en (0, 0, 10).

Solución Sean P la matriz de proyección y D la matriz de datos para S usando coordenadas homogéneas. La matriz de datos para la imagen de S es

$$PD = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/10 & 1 \end{bmatrix} \overset{\text{Vértice:}}{\overset{\begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix}}{\begin{bmatrix} 3 & 5 & 5 & 3 & 3 & 5 & 5 & 3 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 5 & 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}}}$$

$$= \begin{bmatrix} 3 & 5 & 5 & 3 & 3 & 5 & 5 & 3 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .5 & .5 & .5 & .5 & .6 & .6 & .6 & .6 \end{bmatrix}$$

Para obtener las coordenadas en $\mathbb{R}^3$, use (1) y divida las tres entradas superiores de cada columna entre la entrada correspondiente de la cuarta fila:



S bajo la transformación de perspectiva.

$$\overset{\text{Vértice:}}{\overset{\begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix}}{\begin{bmatrix} 6 & 10 & 10 & 6 & 5 & 8.3 & 8.3 & 5 \\ 2 & 2 & 0 & 0 & 1.7 & 1.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}}$$

WEB

El sitio web de este libro contiene algunas aplicaciones interesantes para los gráficos por computadora, incluyendo un análisis más profundo de las proyecciones en perspectiva. Uno de los proyectos para computadora presentados en el sitio web involucra una animación sencilla.

NOTA NUMÉRICA

El movimiento continuo de objetos gráficos tridimensionales requiere cálculos intensivos con matrices de 4×4, sobre todo cuando las superficies se *detallan* para parecer realistas, con textura e iluminación apropiadas. Las *estaciones de trabajo para gráficos* tienen operaciones de matrices de 4×4 y algoritmos de gráficos integrados en sus microprocesadores y circuitos. Dichas estaciones pueden realizar las miles de millones de multiplicaciones por segundo necesarias para presentar la animación realista a color en los programas de juego tridimensionales.[2]

Lectura adicional

James D. Foley, Andries van Dam, Steven K. Feiner y John F. Hughes, *Computer Graphics: Principles and Practice,* 3a. ed., (Boston, MA: Addison-Wesley, 2002), capítulos 5 y 6.

PROBLEMA DE PRÁCTICA

La rotación de una figura alrededor de un punto **p** en $\mathbb{R}^2$ se logra trasladando la figura mediante $-\mathbf{p}$, girándola alrededor del origen, y trasladándola entonces de regreso por medio de **p**. Vea la figura 7. Utilizando coordenadas homogéneas, construya la matriz de 3×3 que gira los puntos en $-30°$ con respecto al punto $(-2, 6)$.

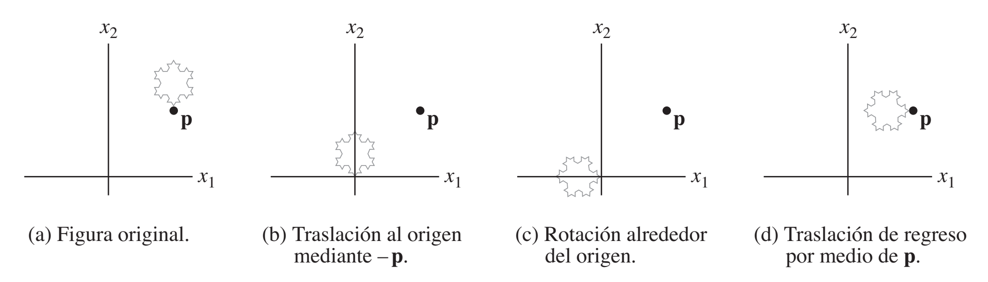


FIGURA 7 Rotación de una figura alrededor del punto **p**.

2.7 EJERCICIOS

1. ¿Qué matriz de 3×3 tendrá el mismo efecto sobre las coordenadas homogéneas para $\mathbb{R}^2$ que el de la matriz de trasquilado A vista en el ejemplo 2?
2. Use la multiplicación de matrices para encontrar la imagen del triángulo con matriz de datos $D = \begin{bmatrix} 5 & 2 & 4 \\ 0 & 2 & 3 \end{bmatrix}$ bajo la transformación que refleja los puntos sobre el eje y. Bosqueje tanto el triángulo original como su imagen.

En los ejercicios 3 a 8, y usando coordenadas homogéneas, encuentre las matrices de 3×3 que producen las transformaciones bidimensionales compuestas descritas.

[2]Vea Jan Ozer, "High-Performance Graphics Boards", *PC Magazine* **19**, 1 de septiembre de 2000, págs. 187-200. También, "The Ultimate Upgrade Guide: Moving On Up", *PC Magazine* **21**, 29 de enero de 2002, págs. 82-91.

3. Trasladar mediante (3, 1), y luego girar en 45° alrededor del origen.

4. Trasladar mediante (−2, 3), y luego escalar la coordenada x por 0.8 y la coordenada y por 1.2.

5. Reflejar puntos a través del eje x, y luego girar en 30° alrededor del origen.

6. Rotar puntos en 30°, y luego reflejarlos a través del eje x.

7. Rotar puntos en 60° alrededor del punto (6, 8).

8. Rotar puntos en 45° alrededor del punto (3, 7).

9. Una matriz de datos D de 2×200 contiene las coordenadas de 200 puntos. Encuentre el número de multiplicaciones que se requieren para transformar estos puntos usando dos matrices arbitrarias A y B de 2×2. Considere las dos posibilidades $A(BD)$ y $(AB)D$. Analice las implicaciones de sus resultados para los cálculos de gráficos por computadora.

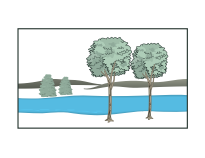
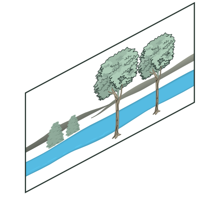
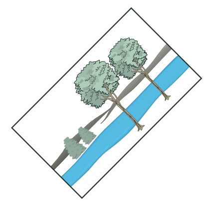
10. Considere las siguientes transformaciones geométricas bidimensionales: D, una dilación (en la que se escalan las coordenadas x y y por el mismo factor); R, una rotación; y T, una traslación. ¿*Conmuta* D con R? Esto es, ¿es $D(R(\mathbf{x})) = R(D(\mathbf{x}))$ para toda $\mathbf{x}$ en $\mathbb{R}^2$? ¿Conmuta D con T? ¿Conmuta R con T?

11. Una rotación en la pantalla de una computadora a veces se implementa como el producto de dos transformaciones de trasquilado y escalamiento, lo cual pueden acelerar los cálculos para determinar cómo aparece realmente una imagen gráfica en términos de los pixeles de la pantalla. (La pantalla consiste en filas y columnas de puntos pequeños, llamados *pixeles*.) La primera transformación A_1 traslada verticalmente y luego comprime cada columna de pixeles; la segunda transformación A_2 traslada horizontalmente y luego expande cada fila de pixeles. Sean

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} \sec\varphi & -\tan\varphi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Muestre que la composición de las dos transformaciones es una rotación en $\mathbb{R}^2$.

12. Una rotación en $\mathbb{R}^2$, por lo general, requiere cuatro multiplicaciones. Encuentre el siguiente producto y muestre que la matriz para una rotación puede factorizarse en tres transformaciones de trasquilado (cada una de las cuales requiere de sólo una multiplicación).

$$\begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ \operatorname{sen}\varphi & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

13. Las transformaciones usuales en coordenadas homogéneas para gráficos por computadora en dos dimensiones usan matrices de 3×3 en la forma

$$\begin{bmatrix} A & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

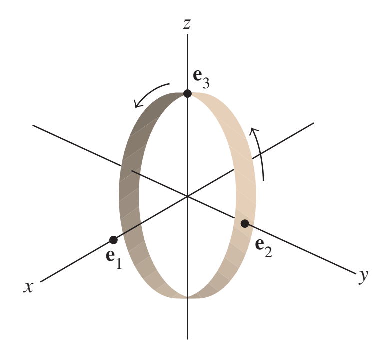
donde A es una matriz de 2×2 y $\mathbf{p}$ está en $\mathbb{R}^2$. Demuestre que una transformación de este tipo equivale a una transformación lineal sobre $\mathbb{R}^2$ seguida por una traslación. [*Sugerencia:* Encuentre una factorización de matrices apropiada en la que intervengan matrices partidas.]

14. Muestre que la transformación del ejercicio 7 es equivalente a una rotación alrededor del origen seguida por una traslación mediante $\mathbf{p}$. Encuentre $\mathbf{p}$.

15. ¿Qué vector en $\mathbb{R}^3$ tiene las coordenadas homogéneas $(\frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, \frac{1}{24})$?

16. ¿Son (1, −2, 3, 4) y (10, −20, 30, 40) coordenadas homogéneas para el mismo punto en $\mathbb{R}^3$? ¿Por qué sí o por qué no?

17. Proporcione la matriz de 4×4 que gira puntos en $\mathbb{R}^3$ alrededor del eje x a través de un ángulo de 60°. (Vea la figura.)



18. Proporcione la matriz de 4×4 que gira puntos en $\mathbb{R}^3$ alrededor del eje z a través de un ángulo de −30°, y luego los traslada mediante $\mathbf{p} = (5, -2, 1)$.

19. Sea S el triángulo con vértices (4.2, 1.2, 4), (6, 4, 2), (2, 2, 6). Encuentre la imagen de S bajo la proyección en perspectiva con centro de proyección en (0, 0, 10).

20. Sea S el triángulo con vértices (9, 3, −5), (12, 8, 2), (1.8, 2.7, 1). Encuentre la imagen de S bajo la proyección en perspectiva con centro de proyección en (0, 0, 10).

Los ejercicios 21 y 22 se refieren a la manera en que se especifica el color para ser mostrado en gráficos por computadora. En una pantalla de computadora, el color se codifica empleando tres números (R, G, B) para indicar la cantidad de energía que un cañón debe transmitir a los puntos fosforescentes rojos, verdes y azules sobre la pantalla de la computadora. (Un cuarto número especifica la luminosidad o intensidad del color.)

21. **[M]** El color real que ve un espectador en una pantalla está influenciado por el tipo específico y la cantidad de material fosforescente que tenga la pantalla. Por ello, cada fabricante de pantallas para computadora debe hacer conversiones entre los datos (R, G, B) y un estándar internacional para color, CIE, el cual usa tres colores primarios llamados X, Y y Z. Una conversión típica para el material fosforescente de persistencia es

$$\begin{bmatrix} .61 & .29 & .150 \\ .35 & .59 & .063 \\ .04 & .12 & .787 \end{bmatrix}\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Un programa de computadora manda un flujo de información sobre el color a la pantalla usando datos del estándar CIE (X, Y, Z). Encuentre la ecuación que convierte estos datos a los datos (R, G, B) que necesita el cañón de electrones de la pantalla.

22. **[M]** La señal transmitida por la televisión comercial describe cada color por medio de un vector (Y, I, Q). Si la pantalla es en blanco y negro, sólo se utiliza la coordenada Y. (Esto proporciona una mejor imagen monocromática que si se usa el estándar CIE para los colores.) La correspondencia entre YIQ y un color "estándar" RGB está dada por

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} .299 & .587 & .114 \\ .596 & -.275 & -.321 \\ .212 & -.528 & .311 \end{bmatrix}\begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

(Un fabricante de pantallas cambiaría las entradas de la matriz para que funcionaran con sus pantallas RGB.) Encuentre la ecuación que convierte los datos YIQ transmitidos por el canal de televisión a los datos RGB necesarios para la pantalla del televisor.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Acomode las matrices de derecha a izquierda para las tres operaciones. Al usar $\mathbf{p} = (-2, 6)$, $\cos(-30°) = \sqrt{3}/2$, y $\text{sen}(-30°) = -0.5$, se tiene:

$$\underset{\text{Trasladar de regreso mediante } p}{\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}}\underset{\text{Girar alrededor del origen}}{\begin{bmatrix} \sqrt{3}/2 & 1/2 & 0 \\ -1/2 & \sqrt{3}/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}}\underset{\text{Trasladar por medio de } -p}{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix}}$$

$$= \begin{bmatrix} \sqrt{3}/2 & 1/2 & \sqrt{3}-5 \\ -1/2 & \sqrt{3}/2 & -3\sqrt{3}+5 \\ 0 & 0 & 1 \end{bmatrix}$$

2.8 SUBESPACIOS DE $\mathbb{R}^n$

Esta sección se enfoca en los importantes conjuntos de vectores en $\mathbb{R}^n$ llamados *subespacios.* Es común que surjan subespacios en relación con alguna matriz A, los cuales proporcionan información útil acerca de la ecuación $A\mathbf{x} = \mathbf{b}$. Los conceptos y la terminología de esta sección se usarán repetidamente a lo largo del resto del libro.[1]

[1] Esta sección se incluye aquí para permitir que los lectores pospongan el estudio de la mayor parte o el total de los siguientes dos capítulos y vayan directamente al capítulo 5, si así lo desean. *Omita* esta sección si planea estudiar el capítulo 4 antes que el capítulo 5.

DEFINICIÓN Un **subespacio** de $\mathbb{R}^n$ es cualquier conjunto H en $\mathbb{R}^n$ que tiene tres propiedades:

a. El vector cero está en H.

b. Para cada $\mathbf{u}$ y $\mathbf{v}$ en H, la suma $\mathbf{u} + \mathbf{v}$ está en H.

c. Para cada $\mathbf{u}$ en H y cada escalar c, el vector $c\mathbf{u}$ está en H.

Dicho con palabras, un subespacio es *cerrado* bajo la suma y la multiplicación escalar. Como se verá en los siguientes ejemplos, casi todos los conjuntos de vectores analizados en el capítulo 1 son subespacios. Por ejemplo, un plano que pasa por el origen es la manera estándar de visualizar el subespacio del ejemplo 1. Vea la figura 1.

FIGURA 1 Gen$\{\mathbf{v}_1, \mathbf{v}_2\}$ tiene un plano a través del origen.

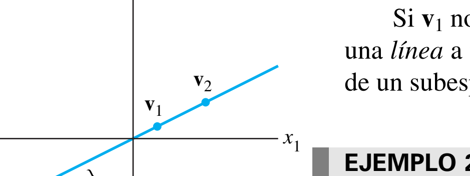
EJEMPLO 1 Si $\mathbf{v}_1$ y $\mathbf{v}_2$ están en $\mathbb{R}^n$ y $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$, entonces H es un subespacio de $\mathbb{R}^n$. Para verificar este enunciado, observe que el vector cero está en H (porque $0\mathbf{v} + 0\mathbf{u}$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$). Ahora tome dos vectores arbitrarios en H, por ejemplo,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 \quad \text{y} \quad \mathbf{v} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

Entonces,

$$\mathbf{u} + \mathbf{v} = (s_1 + t_1)\mathbf{v}_1 + (s_2 + t_2)\mathbf{v}_2$$

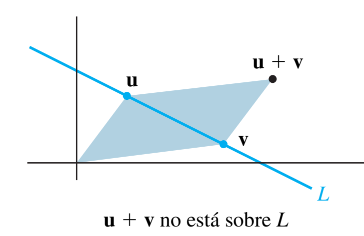
lo cual muestra que $\mathbf{u} + \mathbf{v}$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ y, por lo tanto, está en H. Asimismo, para cualquier escalar c, el vector $c\mathbf{u}$ está en H, porque $c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$.



$\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{v}_2 = k\mathbf{v}_1$

Si $\mathbf{v}_1$ no es cero, y si $\mathbf{v}_2$ es un múltiplo de $\mathbf{v}_1$, entonces $\mathbf{v}_1$ y $\mathbf{v}_2$ simplemente generan una *línea* a través del origen. Por lo tanto, una línea a través del origen es otro ejemplo de un subespacio.

EJEMPLO 2 Una línea L que *no* pasa por el origen *no* es un subespacio, porque no contiene al origen, como se requiere. También, la figura 2 muestra que L no es cerrada bajo la suma ni bajo la multiplicación escalar.



u + **v** no está sobre *L*   2**w** no está sobre *L*

FIGURA 2

EJEMPLO 3 Para $\mathbf{v}_1, \ldots, \mathbf{v}_p$ en $\mathbb{R}^n$, el conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$ es un subespacio de $\mathbb{R}^n$. La verificación de este enunciado es similar al

argumento dado en el ejemplo 1. Ahora es necesario hacer referencia a Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ como **el subespacio generado por** $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

Observe que $\mathbb{R}^n$ es un subespacio de sí mismo porque tiene las tres propiedades requeridas para un subespacio. Otro subespacio especial es el conjunto que consta exclusivamente del vector cero en $\mathbb{R}^n$. Este conjunto, llamado **subespacio cero**, también satisface las condiciones que demanda un subespacio.

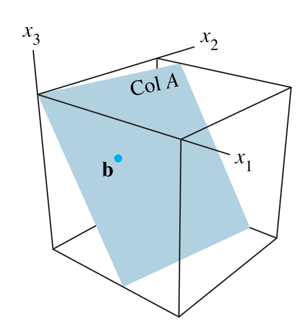
Espacio columna y espacio nulo de una matriz

Los subespacios de $\mathbb{R}^n$ suelen aparecer en las aplicaciones y la teoría en una de dos formas. En ambos casos, es posible relacionar el subespacio con una matriz.

DEFINICIÓN El **espacio columna** de una matriz A es el conjunto Col A de todas las combinaciones lineales de las columnas de A.

Si $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$, con las columnas en $\mathbb{R}^m$, entonces Col A es lo mismo que Gen$\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$. En el ejemplo 3 se muestra que **el espacio columna de una matriz** $\boldsymbol{m \times n}$ **es un subespacio de** $\mathbb{R}^m$.

EJEMPLO 4 Sea $A = \begin{bmatrix} 1 & -3 & -4 \\ -4 & 6 & -2 \\ -3 & 7 & 6 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 3 \\ 3 \\ -4 \end{bmatrix}$. Determine si $\mathbf{b}$ está en el espacio columna de A.



Solución El vector $\mathbf{b}$ es una combinación lineal de las columnas de A si, y sólo si, $\mathbf{b}$ puede escribirse como $A\mathbf{x}$ para alguna $\mathbf{x}$, esto es, si, y sólo si, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución. Al reducir por filas la matriz aumentada $[A \ \ \mathbf{b}]$, se tiene

$$\begin{bmatrix} 1 & -3 & -4 & 3 \\ -4 & 6 & -2 & 3 \\ -3 & 7 & 6 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & -4 & 3 \\ 0 & -6 & -18 & 15 \\ 0 & -2 & -6 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & -4 & 3 \\ 0 & -6 & -18 & 15 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

se concluye que $A\mathbf{x} = \mathbf{b}$ es consistente, y $\mathbf{b}$ está en Col A.

La solución del ejemplo 4 muestra que, cuando un sistema de ecuaciones lineales está escrito en la forma $A\mathbf{x} = \mathbf{b}$, el espacio columna de A es el conjunto de todas las $\mathbf{b}$ para las cuales el sistema tiene una solución.

DEFINICIÓN El **espacio nulo** de una matriz A es el conjunto Nul A de todas las soluciones posibles para la ecuación homogénea $A\mathbf{x} = \mathbf{0}$.

Cuando A tiene n columnas, las soluciones de $A\mathbf{x} = \mathbf{0}$ pertenecen a $\mathbb{R}^n$, y el espacio nulo de A es un subconjunto de $\mathbb{R}^n$. De hecho, Nul A tiene las propiedades de un *subespacio* de $\mathbb{R}^n$.

TEOREMA 12

El espacio nulo de una matriz A de $m \times n$ es un subespacio de $\mathbb{R}^n$. De manera equivalente, el conjunto de todas las soluciones posibles para un sistema $A\mathbf{x} = \mathbf{0}$ de m ecuaciones lineales homogéneas con n incógnitas es un subespacio de $\mathbb{R}^n$.

DEMOSTRACIÓN El vector cero está en Nul A (porque $A\mathbf{0} = \mathbf{0}$). Para demostrar que Nul A satisface las otras dos propiedades que se requieren para conformar un subespacio, se toman cualesquiera $\mathbf{u}$ y $\mathbf{v}$ en Nul A. Esto es, se supone que $A\mathbf{u} = \mathbf{0}$ y $A\mathbf{v} = \mathbf{0}$. Luego, por una propiedad de la multiplicación de matrices,

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Entonces $\mathbf{u} + \mathbf{v}$ satisface $A\mathbf{x} = \mathbf{0}$, y así $\mathbf{u} + \mathbf{v}$ está en Nul A. De igual modo, para cualquier escalar c, $A(c\mathbf{u}) = c(A\mathbf{u}) = c(\mathbf{0}) = \mathbf{0}$, lo cual muestra que $c\mathbf{u}$ está en Nul A. ■

Para probar si un vector dado $\mathbf{v}$ está en Nul A, simplemente se calcula $A\mathbf{v}$ para ver si $A\mathbf{v}$ es el vector cero. Como Nul A se describe por medio de una condición que debe comprobarse para cada vector, se dice que el espacio nulo está definido *implícitamente*. En contraste, el espacio columna se define *explícitamente*, porque los vectores en Col A se pueden construir (por medio de combinaciones lineales) a partir de las columnas de A. Para crear una descripción explícita de Nul A, resuelva la ecuación $A\mathbf{x} = \mathbf{0}$ y escriba la solución en forma vectorial paramétrica. (Vea el ejemplo 6 que se presenta en la siguiente página.)[2]

Bases para un subespacio

Dado que, por lo general, un subespacio contiene un número infinito de vectores, algunos problemas relacionados con subespacios se manejan mejor trabajando con un conjunto finito y pequeño de vectores que genera el subespacio. Entre más pequeño sea el conjunto, mejor. Es posible demostrar que el conjunto generador más pequeño debe ser linealmente independiente.

DEFINICIÓN

Una **base** para un subespacio H de $\mathbb{R}^n$ es un conjunto linealmente independiente en H que genera H.

EJEMPLO 5 Las columnas de una matriz invertible $n \times n$ forman una base para todo $\mathbb{R}^n$ porque son linealmente independientes y generan $\mathbb{R}^n$, de acuerdo con el teorema de la matriz invertible. Una matriz de este tipo es la matriz identidad $n \times n$. Sus columnas se denotan mediante $\mathbf{e}_1, \ldots, \mathbf{e}_n$:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

El conjunto $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ es la **base estándar** para $\mathbb{R}^n$. Vea la figura 3. ■

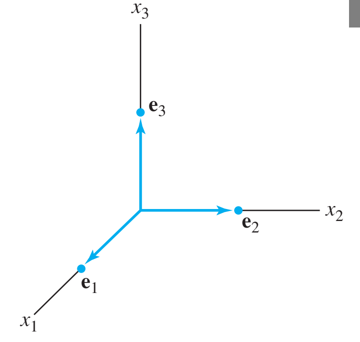


FIGURA 3
La base estándar para $\mathbb{R}^3$.

[2] El contraste entre Nul A y Col A se analiza con mayor amplitud en la sección 4.2.

En el ejemplo siguiente se muestra que el procedimiento estándar para escribir el conjunto solución de $A\mathbf{x} = \mathbf{0}$ en la forma vectorial paramétrica en realidad identifica una base para Nul A. Este hecho se utilizará durante todo el capítulo 5.

EJEMPLO 6 Encuentre una base para el espacio nulo de la matriz

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

Solución Primero, escriba la solución de $A\mathbf{x} = \mathbf{0}$ en forma vectorial paramétrica:

$$A \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 - 2x_2 \qquad - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

La solución general es $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, con x_2, x_4 y x_5 libres.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\substack{\uparrow \\ \mathbf{w}}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}}$$

$$= x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w} \tag{1}$$

La ecuación (1) muestra que Nul A coincide con el conjunto de todas las combinaciones lineales de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$. Esto es, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ genera Nul A. De hecho, esta construcción de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ las vuelve, de manera automática, linealmente independientes, porque (1) muestra que $\mathbf{0} = x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$ solamente si los pesos x_2, x_4 y x_5 son todos cero. (Examine las entradas 2, 4 y 5 del vector $x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$.) Por lo tanto, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es una *base* para Nul A.

La determinación de una base para el espacio columna de una matriz representa, de hecho, menos trabajo que encontrar una base para el espacio nulo. Sin embargo, el método necesita cierta explicación. Enseguida se presenta un caso sencillo.

EJEMPLO 7 Encuentre una base para el espacio columna de la matriz

$$B = \begin{bmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Solución Denote las columnas de B mediante $\mathbf{b}_1, \ldots, \mathbf{b}_5$ y observe que $\mathbf{b}_3 = -3\mathbf{b}_1 + 2\mathbf{b}_2$, y que $\mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2$. El que $\mathbf{b}_3$ y $\mathbf{b}_4$ sean combinaciones de las columnas pivote implica que cualquier combinación de $\mathbf{b}_1, \ldots, \mathbf{b}_5$ es en realidad sólo una combinación

de $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_5$. En efecto, si $\mathbf{v}$ es cualquier vector en Col B, por ejemplo,

$$\mathbf{v} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + c_4\mathbf{b}_4 + c_5\mathbf{b}_5$$

entonces, sustituyendo $\mathbf{b}_3$ y $\mathbf{b}_4$, se puede escribir $\mathbf{v}$ en la forma

$$\mathbf{v} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3(-3\mathbf{b}_1 + 2\mathbf{b}_2) + c_4(5\mathbf{b}_1 - \mathbf{b}_2) + c_5\mathbf{b}_5$$

lo cual es una combinación lineal de $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_5$. Así que $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$ genera Col B. También, $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_5$ son linealmente independientes, porque son columnas de una matriz identidad. Por lo tanto, las columnas pivote de B forman una base para Col B. ■

La matriz B del ejemplo 7 estaba en forma escalonada reducida. Para manejar una matriz general A, recuerde que las relaciones de dependencia lineal entre las columnas de A se pueden expresar en la forma $A\mathbf{x} = \mathbf{0}$ para alguna $\mathbf{x}$. (Si algunas columnas no intervienen en una relación de dependencia dada, entonces las entradas correspondientes de $\mathbf{x}$ son cero.) Cuando A se reduce por filas a la forma escalonada B, las columnas se cambian drásticamente, pero las ecuaciones $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ tienen el mismo conjunto de soluciones. Esto es, las columnas de A tienen *exactamente las mismas relaciones de dependencia lineal* que las columnas de B.

EJEMPLO 8 Se puede verificar que la matriz

$$A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_5\,] = \begin{bmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{bmatrix}$$

es equivalente por filas a la matriz B del ejemplo 7. Encuentre una base para Col A.

Solución A partir del ejemplo 7, las columnas pivote de A son las columnas 1, 2 y 5. También, $\mathbf{b}_3 = -\,3\mathbf{b}_1 + 2\mathbf{b}_2$ y $\mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2$. Como las operaciones por fila no afectan las relaciones de dependencia lineal entre las columnas de la matriz, se debe tener que

$$\mathbf{a}_3 = -3\mathbf{a}_1 + 2\mathbf{a}_2 \quad \text{y} \quad \mathbf{a}_4 = 5\mathbf{a}_1 - \mathbf{a}_2$$

Compruebe que esto sea cierto. Por el argumento del ejemplo 7, $\mathbf{a}_3$ y $\mathbf{a}_4$ no se necesitan para generar el espacio columna de A. También, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ debe ser linealmente independiente, porque cualquier relación de dependencia entre $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_5$ implicaría la misma relación de dependencia entre $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_5$. Como $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$ es linealmente independiente, también $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ lo es y, por lo tanto, es una base para Col A. ■

El argumento del ejemplo 8 se puede adaptar para demostrar el teorema siguiente.

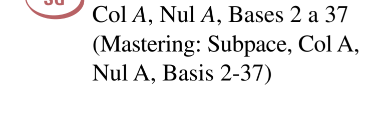
TEOREMA 13 Las columnas pivote de una matriz A forman una base para el espacio columna de A.

Advertencia: Tenga cuidado de utilizar *las columnas pivote de la misma* A para la base de Col A. Las columnas de una forma escalonada B a menudo no están en el espacio columna de A. (Como puede observarse en los ejemplos 7 y 8, todas las columnas de B tienen ceros en sus últimas entradas y no pueden generar las columnas de A.)

PROBLEMAS DE PRÁCTICA

1. Sea $A = \begin{bmatrix} 1 & -1 & 5 \\ 2 & 0 & 7 \\ -3 & -5 & -3 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} -7 \\ 3 \\ 2 \end{bmatrix}$. ¿Está $\mathbf{u}$ en Nul A? ¿Está $\mathbf{u}$ en Col A? Justifique sus respuestas.

2. Dada $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, encuentre un vector en Nul A y un vector en Col A.



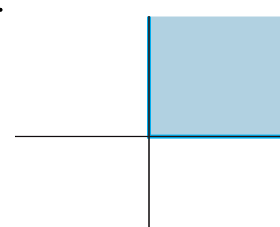
Dominio de subespacios, Col A, Nul A, Bases 2 a 37 (Mastering: Subpace, Col A, Nul A, Basis 2-37)

3. Suponga que una matriz A de $n \times n$ es invertible. ¿Qué puede decirse acerca de Col A? ¿Qué acerca de Nul A?

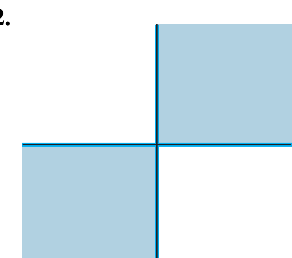
2.8 EJERCICIOS

En los ejercicios 1 a 4 se muestran conjuntos en $\mathbb{R}^2$. Suponga que los conjuntos incluyen las líneas de frontera. En cada caso, proporcione una razón específica por la cual el conjunto H *no* es un subespacio de $\mathbb{R}^2$. (Por ejemplo, encuentre dos vectores en H cuya suma *no* esté en H, o encuentre un vector en H con un múltiplo escalar que no esté en H. Trace un esquema.)

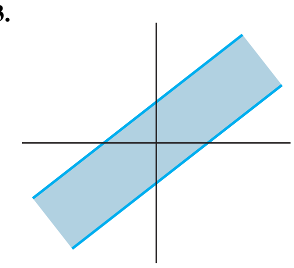
1.



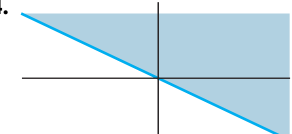
2.



3.



4.



5. Sean $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \\ -5 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ -5 \\ 8 \end{bmatrix}$, y $\mathbf{w} = \begin{bmatrix} 8 \\ 2 \\ -9 \end{bmatrix}$. Determine si $\mathbf{w}$ está en el subespacio de $\mathbb{R}^3$ generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

6. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 4 \\ -7 \\ 9 \\ 7 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 5 \\ -8 \\ 6 \\ 5 \end{bmatrix}$, y $\mathbf{u} = \begin{bmatrix} -4 \\ 10 \\ -7 \\ -5 \end{bmatrix}$. Determine si $\mathbf{u}$ está en el subespacio de $\mathbb{R}^4$ generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

7. Sean $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -8 \\ 6 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -3 \\ 8 \\ -7 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} -4 \\ 6 \\ -7 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 6 \\ -10 \\ 11 \end{bmatrix}$, y $A = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$.

a. ¿Cuántos vectores hay en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

b. ¿Cuántos vectores hay en Col A?

c. ¿Está $\mathbf{p}$ en Col A? ¿Por qué sí o por qué no?

8. Sean $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ -6 \\ 3 \end{bmatrix}$, y $\mathbf{p} = \begin{bmatrix} 1 \\ 14 \\ -9 \end{bmatrix}$. Determine si $\mathbf{p}$ está en Col A, donde $A = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$.

9. Con A y $\mathbf{p}$ como en el ejercicio 7, determine si $\mathbf{p}$ está en Nul A.

10. Con $\mathbf{u} = (-2, 3, 1)$ y A como en ejercicio 8, determine si $\mathbf{u}$ está en Nul A.

En los ejercicios 11 y 12, proporcione enteros p y q tales que Nul A sea un subespacio de $\mathbb{R}^p$ y Col A un subespacio de $\mathbb{R}^q$.

11. $A = \begin{bmatrix} 3 & 2 & 1 & -5 \\ -9 & -4 & 1 & 7 \\ 9 & 2 & -5 & 1 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ -5 & -1 & 0 \\ 2 & 7 & 11 \end{bmatrix}$

13. Para A como en el ejercicio 11, encuentre un vector diferente de cero en Nul A y un vector diferente de cero en Col A.

14. Para A como en el ejercicio 12, encuentre un vector diferente de cero en Nul A y un vector diferente de cero en Col A.

Determine cuáles conjuntos de los ejercicios 15 a 20 son bases para $\mathbb{R}^2$ y $\mathbb{R}^3$. Justifique sus respuestas.

15. $\begin{bmatrix} 5 \\ -2 \end{bmatrix}, \begin{bmatrix} 10 \\ -3 \end{bmatrix}$

16. $\begin{bmatrix} -4 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

17. $\begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \\ 5 \end{bmatrix}$

18. $\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -5 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ -5 \end{bmatrix}$

19. $\begin{bmatrix} 3 \\ -8 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \\ -5 \end{bmatrix}$

20. $\begin{bmatrix} 1 \\ -6 \\ -7 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ 7 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 8 \\ 9 \end{bmatrix}$

En los ejercicios 21 y 22 señale cada enunciado como verdadero o falso. Justifique sus respuestas.

21. a. Un subespacio de $\mathbb{R}^n$ es cualquier conjunto H tal que (i) el vector cero está en H, (ii) $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} + \mathbf{v}$ están en H, y (iii) c es un escalar y $c\mathbf{u}$ está en H.

b. Si $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en $\mathbb{R}^n$, entonces Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es lo mismo que el espacio columna de la matriz $[\mathbf{v}_1 \ \cdots \ \mathbf{v}_p]$.

c. El conjunto de todas las soluciones de un sistema de m ecuaciones homogéneas en n incógnitas es un subespacio de $\mathbb{R}^m$.

d. Las columnas de una matriz invertible $n \times n$ forman una base para $\mathbb{R}^n$.

e. Las operaciones de fila no afectan las relaciones de dependencia lineal entre las columnas de una matriz.

22. a. Un subconjunto H de $\mathbb{R}^n$ es un subespacio si el vector cero está en H.

b. Dados los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_p$ en $\mathbb{R}^n$, el conjunto de todas las combinaciones lineales de estos vectores es un subespacio de $\mathbb{R}^n$.

c. El espacio nulo de una matriz $m \times n$ es un subespacio de $\mathbb{R}^n$.

d. El espacio columna de una matriz A es el conjunto de soluciones de $A\mathbf{x} = \mathbf{b}$.

e. Si B es una forma escalonada de una matriz A, entonces las columnas pivote de B forman una base para Col A.

En los ejercicios 23 a 26 se presenta una matriz A y una forma escalonada de A. Encuentre una base para Col A y una base para Nul A.

23. $A = \begin{bmatrix} 4 & 5 & 9 & -2 \\ 6 & 5 & 1 & 12 \\ 3 & 4 & 8 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 6 & -5 \\ 0 & 1 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

24. $A = \begin{bmatrix} -3 & 9 & -2 & -7 \\ 2 & -6 & 4 & 8 \\ 3 & -9 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 6 & 9 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

25. $A = \begin{bmatrix} 1 & 4 & 8 & -3 & -7 \\ -1 & 2 & 7 & 3 & 4 \\ -2 & 2 & 9 & 5 & 5 \\ 3 & 6 & 9 & -5 & -2 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & 4 & 8 & 0 & 5 \\ 0 & 2 & 5 & 0 & -1 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

26. $A = \begin{bmatrix} 3 & -1 & 7 & 3 & 9 \\ -2 & 2 & -2 & 7 & 5 \\ -5 & 9 & 3 & 3 & 4 \\ -2 & 6 & 6 & 3 & 7 \end{bmatrix}$

$\sim \begin{bmatrix} 3 & -1 & 7 & 0 & 6 \\ 0 & 2 & 4 & 0 & 3 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

27. Construya una matriz A de 3×3 y un vector $\mathbf{b}$ distinto de cero en forma tal que $\mathbf{b}$ esté en Col A, pero $\mathbf{b}$ no sea lo mismo que alguna de las columnas de A.

28. Construya una matriz A de 3×3 y un vector $\mathbf{b}$ en forma tal que $\mathbf{b}$ *no* esté en Col A.

29. Construya una matriz A de 3×3 distinta de cero y un vector $\mathbf{b}$ diferente de cero en forma tal que $\mathbf{b}$ esté en Nul A.

30. Suponga que las columnas de una matriz $A = [\mathbf{a}_1 \cdots \mathbf{a}_p]$ son linealmente independientes. Explique por qué $\{\mathbf{a}_1, \ldots, \mathbf{a}_p\}$ es una base para Col A.

En los ejercicios 31 a 36, responda de manera tan comprensible como sea posible y justifique sus respuestas.

31. Suponga que F es una matriz de 5×5 cuyo espacio columna no es igual a $\mathbb{R}^5$. ¿Qué puede decirse acerca de Nul F?

32. Si R es una matriz de 6×6 y Nul R *no* es el subespacio cero, ¿qué puede decirse acerca de Col R?

33. Si Q es una matriz de 4×4 y Col $Q = \mathbb{R}^4$, ¿qué puede decirse acerca de las soluciones a ecuaciones de la forma $Q\mathbf{x} = \mathbf{b}$ para $\mathbf{b}$ en $\mathbb{R}^4$.

34. Si P es una matriz de 5×5 y Nul P es el subespacio cero, ¿qué puede decirse acerca de las soluciones a ecuaciones de la forma $P\mathbf{x} = \mathbf{b}$ para $\mathbf{b}$ en $\mathbb{R}^5$?

35. ¿Qué puede decirse acerca de Nul B cuando B es una matriz de 5×4 con columnas linealmente independientes?

36. ¿Qué puede decirse acerca de la forma de una matriz A de $m \times n$ cuando las columnas de A constituyen una base para $\mathbb{R}^m$?

[M] En los ejercicios 37 y 38, construya bases para el espacio columna y para el espacio nulo de la matriz A dada. Justifique el trabajo realizado.

37. $A = \begin{bmatrix} 3 & -5 & 0 & -1 & 3 \\ -7 & 9 & -4 & 9 & -11 \\ -5 & 7 & -2 & 5 & -7 \\ 3 & -7 & -3 & 4 & 0 \end{bmatrix}$

38. $A = \begin{bmatrix} 5 & 2 & 0 & -8 & -8 \\ 4 & 1 & 2 & -8 & -9 \\ 5 & 1 & 3 & 5 & 19 \\ -8 & -5 & 6 & 8 & 5 \end{bmatrix}$

CD Espacio columna y espacio nulo (Column Space and Null Space)

CD Una base para Col A (A Basis for Col A)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Para determinar si $\mathbf{u}$ está en Nul A, simplemente calcule

$$A\mathbf{u} = \begin{bmatrix} 1 & -1 & 5 \\ 2 & 0 & 7 \\ -3 & -5 & -3 \end{bmatrix} \begin{bmatrix} -7 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

El resultado muestra que $\mathbf{u}$ está en Nul A. Para decidir si $\mathbf{u}$ está en Col A se requiere más trabajo. Reduzca la matriz aumentada $[A \quad \mathbf{u}]$ a la forma escalonada para determinar si la ecuación $A\mathbf{x} = \mathbf{u}$ es consistente:

$$\begin{bmatrix} 1 & -1 & 5 & -7 \\ 2 & 0 & 7 & 3 \\ -3 & -5 & -3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & -8 & 12 & -19 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & 0 & 0 & 49 \end{bmatrix}$$

La ecuación $A\mathbf{x} = \mathbf{u}$ no tiene solución, entonces $\mathbf{u}$ no está en Col A.

2. En contraste con el problema de práctica 1, encontrar un vector en Nul A requiere más trabajo que probar si un vector específico está en Nul A. Sin embargo, como A ya está en forma escalonada reducida, la ecuación $A\mathbf{x} = \mathbf{0}$ muestra que si $\mathbf{x} = (x_1, x_2, x_3)$, entonces $x_2 = 0$, $x_3 = 0$, y x_1 es una variable libre. Por lo tanto, una base para Nul A es $\mathbf{v} = (1, 0, 0)$. Encontrar sólo un vector en Col A resulta trivial, puesto que cada columna de A está en Col A. En este caso particular, el mismo vector $\mathbf{v}$ se encuentra *tanto en* Nul A *como en* Col A. Para la mayoría de las matrices $n \times n$, el vector cero de $\mathbb{R}^n$ es el único vector que se encuentra tanto en Nul A como en Col A.

3. Si A es invertible, entonces las columnas de A generan $\mathbb{R}^n$, según el teorema de la matriz invertible. Por definición, las columnas de cualquier matriz siempre generan el espacio columna, entonces, en este caso, Col A es todo $\mathbb{R}^n$. En forma simbólica, Col $A = \mathbb{R}^n$. También, como A es invertible, la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial. Esto significa que Nul A es el subespacio cero. En forma simbólica, Nul $A = \{\mathbf{0}\}$.

2.9 DIMENSIÓN Y RANGO

En esta sección se continúa el análisis de los subespacios y las bases para subespacios, iniciando con el concepto de un sistema coordenado. La definición y el ejemplo presentados a continuación pretenden que un término nuevo y útil, *dimensión*, parezca bastante natural, al menos para los subespacios de $\mathbb{R}^3$.

Sistemas de coordenadas

La razón principal para seleccionar la base de un subespacio H, en lugar de simplemente un conjunto generador, es que cada vector de H se puede escribir sólo de una manera como combinación lineal de los vectores de la base. Para ver por qué, suponga que $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ es una base de H, y que un vector $\mathbf{x}$ en H puede generarse de dos maneras, por ejemplo,

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p \quad \text{y} \quad \mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_p\mathbf{b}_p \tag{1}$$

Después, restando se obtiene

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_p - d_p)\mathbf{b}_p \tag{2}$$

Como $\mathcal{B}$ es linealmente independiente, los pesos en (2) deben ser todos cero. Esto es, $c_j = d_j$ para $1 \leq j \leq p$, lo cual muestra que las dos representaciones en (2) son, de hecho, la misma representación.

DEFINICIÓN

Suponga que el conjunto $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p]$ es la base de un subespacio H. Para cada $\mathbf{x}$ en H, las **coordenadas de x relativas a la base** $\mathcal{B}$ son los pesos $c_1, \ldots, c_p$ tales que $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p$, y el vector en $\mathbb{R}^p$

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

se llama **vector de coordenadas de x (relativo a** $\mathcal{B}$**)** o **vector de** $\mathcal{B}$**-coordenadas de x**.[1]

EJEMPLO 1 Sea $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$, y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Entonces $\mathcal{B}$ es una base de $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$ porque $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes. Determine si $\mathbf{x}$ está en H y, si lo está, encuentre el vector de coordenadas de $\mathbf{x}$ relativo a $\mathcal{B}$.

Solución Si $\mathbf{x}$ está en H, entonces la siguiente ecuación vectorial es consistente:

$$c_1\begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

[1]Es importante que los elementos de $\mathcal{B}$ estén numerados porque las entradas de $[\mathbf{x}]_{\mathcal{B}}$ dependen del orden de los vectores en $\mathcal{B}$.

Los escalares c_1, c_2, si existen, son las $\mathcal{B}$-coordenadas de $\mathbf{x}$. Al aplicar operaciones por fila, se tiene que

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Entonces $c_1 = 2$, $c_2 = 3$, y $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. La base $\mathcal{B}$ determina un “sistema de coordenadas” en H, lo cual puede visualizarse por medio de la red mostrada en la figura 1.

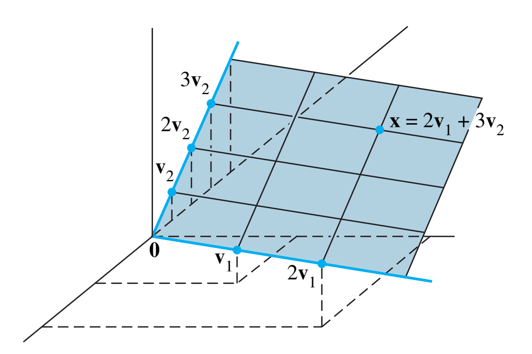


FIGURA 1 Un sistema de coordenadas sobre un plano H en $\mathbb{R}^3$.

Observe que a pesar de que los puntos de H también se encuentran en $\mathbb{R}^3$, están completamente determinados por sus vectores de coordenadas, los cuales pertenecen a $\mathbb{R}^2$. La malla mostrada en el plano de la figura 1 hace que H “se vea” como $\mathbb{R}^2$. La correspondencia $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es una correspondencia uno a uno entre H y $\mathbb{R}^2$ que conserva las combinaciones lineales. A una correspondencia de este tipo se le llama *isomorfismo*, y se dice que H es *isomorfo* a $\mathbb{R}^2$.

En general, si $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ es una base para H, entonces la función $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es una correspondencia uno a uno que permite a H verse y funcionar igual que $\mathbb{R}^p$ (aunque los propios vectores de H puedan tener más de p entradas). (En la sección 4.4 se presentan más detalles.)

La dimensión de un subespacio

Se puede demostrar que si un subespacio H tiene una base de p vectores, entonces cualquier base de H debe consistir en exactamente p vectores. (Vea los ejercicios 27 y 28.) Por lo tanto, la siguiente definición tiene sentido.

DEFINICIÓN La **dimensión** de un subespacio H diferente de cero, denotada mediante dim H, es el número de vectores que hay en cualquier base de H. La dimensión del subespacio cero $\{\mathbf{0}\}$ es, por definición, cero.[2]

[2]El subespacio cero *no* tiene base (porque el vector cero forma, por sí mismo, un conjunto linealmente dependiente).

El espacio $\mathbb{R}^n$ tiene dimensión n. Cada base para $\mathbb{R}^n$ consiste en n vectores. Un plano a través de $\mathbf{0}$ en $\mathbb{R}^3$ es bidimensional, y una línea a través de $\mathbf{0}$ es unidimensional.

EJEMPLO 2 Recuerde que el espacio nulo de la matriz A vista en el ejemplo 6, sección 2.8, tenía una base de tres vectores. Así que la dimensión de Nul A en este caso es 3. Observe cómo cada vector de base corresponde a una variable libre en la ecuación $A\mathbf{x} = \mathbf{0}$. La construcción realizada aquí siempre produce una base de este modo. Entonces, para encontrar la dimensión de Nul A, basta con identificar y contar el número de variables libres en $A\mathbf{x} = \mathbf{0}$.

DEFINICIÓN El **rango** de una matriz A, denotado mediante rango A, es la dimensión del espacio columna de A.

Como las columnas pivote de A forman una base para Col A, el rango de A es simplemente el número de columnas pivote en A.

EJEMPLO 3 Determine el rango de la matriz

$$A = \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 4 & 7 & -4 & -3 & 9 \\ 6 & 9 & -5 & 2 & 4 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix}$$

Solución Reduzca A a la forma escalonada:

$$A \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & -6 & 4 & 14 & -20 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Columnas pivote

La matriz A tiene tres columnas pivote, así que rango $A = 3$.

La reducción por filas del ejemplo 3 revela que hay dos variables libres en $A\mathbf{x} = \mathbf{0}$, porque dos de las cinco columnas de A *no* son columnas pivote. (Las columnas que no son pivote corresponden a las variables libres de $A\mathbf{x} = \mathbf{0}$.) Como el número de columnas pivote más el número de columnas que no son pivote es exactamente el número de columnas, las dimensiones de Col A y Nul A tienen la siguiente conexión útil. (Si desea ver detalles adicionales, consulte el teorema de rango presentado en la sección 4.6.)

TEOREMA 14 **El teorema de rango**

Si una matriz A tiene n columnas, entonces rango A + dim Nul $A = n$.

El teorema siguiente es importante para las aplicaciones y se necesitará en los capítulos 5 y 6. El teorema (demostrado en la sección 4.5) evidentemente es verosímil, si

se piensa en un subespacio p-dimensional como isomorfo a $\mathbb{R}^p$. El teorema de la matriz invertible muestra que p vectores de $\mathbb{R}^p$ son linealmente independientes si, y sólo si, también generan $\mathbb{R}^p$.

TEOREMA 15

El teorema de la base

Sea H un subespacio p-dimensional de $\mathbb{R}^n$. Cualquier conjunto linealmente independiente de exactamente p elementos en H automáticamente es una base de H. También, cualquier conjunto de p elementos de H que genere H es automáticamente una base para H.

Rango y el teorema de la matriz invertible

Los diversos conceptos de espacio vectorial asociados con una matriz proporcionan varios enunciados más para el teorema de la matriz invertible. Estos enunciados se presentan enseguida como una continuación del teorema original presentado en la sección 2.3.

TEOREMA

El teorema de la matriz invertible (continuación)

Sea A una matriz $n \times n$. Entonces, cada uno de los siguientes enunciados es equivalente al enunciado de que A es una matriz invertible.

m. Las columnas de A forman una base de $\mathbb{R}^n$.
n. $\text{Col } A = \mathbb{R}^n$.
o. $\dim \text{Col } A = n$.
p. $\text{rango } A = n$.
q. $\text{Nul } A = \{\mathbf{0}\}$.
r. $\dim \text{Nul } A = 0$.

DEMOSTRACIÓN El enunciado (m) es, por lógica, equivalente a los enunciados (e) y (h) relativos a la independencia lineal y a la generación. Los otros cinco enunciados se vinculan a los primeros del teorema por medio de la siguiente cadena de implicaciones casi triviales.

$$(g) \Rightarrow (n) \Rightarrow (o) \Rightarrow (p) \Rightarrow (r) \Rightarrow (q) \Rightarrow (d)$$

El enunciado (g), el cual indica que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para cada $\mathbf{b}$ en $\mathbb{R}^n$, implica a (n), porque Col A es precisamente el conjunto de todas las $\mathbf{b}$ tales que la ecuación $A\mathbf{x} = \mathbf{b}$ sea consistente. Las implicaciones (n) $\Rightarrow$ (o) $\Rightarrow$ (p) se siguen de las definiciones de *dimensión* y *rango*. Si el rango de A es n, el número de columnas de A, entonces $\dim \text{Nul } A = 0$, por el teorema del rango, y así $\text{Nul } A = \{\mathbf{0}\}$. De modo que (p) $\Rightarrow$ (r) $\Rightarrow$ (q). Asimismo, (q) implica que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial, que es el enunciado (d). Dado que los enunciados (d) y (g) ya son conocidos como equivalentes al enunciado de que A es invertible, la comprobación está completa. ■

SG Tabla expandida para el TMI 2 a 39 (Expanded Table for the IMT 2-39)

NOTAS NUMÉRICAS

Muchos de los algoritmos analizados en este texto resultan útiles para entender conceptos y realizar manualmente cálculos sencillos. Sin embargo, a menudo los algoritmos no son aplicables a problemas de gran escala en la vida real.

Un buen ejemplo de lo anterior es la determinación del rango. Podría parecer sencillo reducir una matriz a su forma escalonada y contar los pivotes. Pero aunque se realicen cálculos precisos sobre una matriz cuyas entradas estén especificadas exactamente, las operaciones de fila pueden cambiar el rango aparente de una matriz. Por ejemplo, si el valor de x en la matriz $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ no se almacena exactamente como 7 en una computadora, entonces el rango puede ser 1 o 2, dependiendo de si la computadora considera a $x - 7$ como cero.

CD El comando `rank` (The `rank` command)

En las aplicaciones prácticas, es frecuente que el rango efectivo de una matriz A se determine a partir de la descomposición del valor singular de A, el cual se estudiará en la sección 7.4.

PROBLEMAS DE PRÁCTICA

1. Determine la dimensión del subespacio H de $\mathbb{R}^3$ generado por los vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. (Primero encuentre una base para H.)

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -8 \\ 6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -7 \\ -1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 6 \\ -7 \end{bmatrix}$$

2. Considere la base $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ .2 \end{bmatrix}, \begin{bmatrix} .2 \\ 1 \end{bmatrix} \right\}$ para $\mathbb{R}^2$. Si $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, ¿qué es $\mathbf{x}$?

3. ¿Podría $\mathbb{R}^3$ contener a un subespacio cuatridimensional? Explique su respuesta.

2.9 EJERCICIOS

En los ejercicios 1 y 2, encuentre el vector $\mathbf{x}$ determinado por el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ dado y la base $\mathcal{B}$ dada. Ilustre cada respuesta con una figura, como en la solución al problema de práctica 2.

1. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

2. $\mathcal{B} = \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$

En los ejercicios 3 a 6, el vector $\mathbf{x}$ está en un subespacio H que tiene una base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Encuentre el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$.

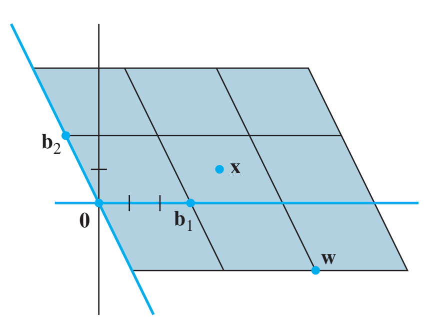
3. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -2 \\ 7 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$

4. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} -7 \\ 5 \end{bmatrix}$

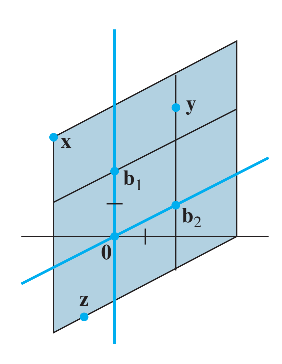
5. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ -7 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 10 \\ -7 \end{bmatrix}$

6. $\mathbf{b}_1 = \begin{bmatrix} -3 \\ 1 \\ -4 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 7 \\ 5 \\ -6 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 11 \\ 0 \\ 7 \end{bmatrix}$

7. Sea $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Use la figura para estimar $[\mathbf{w}]_{\mathcal{B}}$ y $[\mathbf{x}]_{\mathcal{B}}$. Confirme su estimación de $[\mathbf{x}]_{\mathcal{B}}$ usándola junto con $\{\mathbf{b}_1, \mathbf{b}_2\}$ para calcular $\mathbf{x}$.



8. Sean $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} -1 \\ -2.5 \end{bmatrix}$, y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Use la figura para estimar $[\mathbf{x}]_{\mathcal{B}}$, $[\mathbf{y}]_{\mathcal{B}}$ y $[\mathbf{z}]_{\mathcal{B}}$. Confirme su estimación de $[\mathbf{y}]_{\mathcal{B}}$ y $[\mathbf{z}]_{\mathcal{B}}$ usándolas junto con $\{\mathbf{b}_1, \mathbf{b}_2\}$ para calcular $\mathbf{y}$ y $\mathbf{z}$.



En los ejercicios 9 a 12 se presentan una matriz A y una forma escalonada de A. Encuentre bases para Col A y Nul A, y luego establezca las dimensiones de estos subespacios.

9. $A = \begin{bmatrix} 1 & -3 & 2 & -4 \\ -3 & 9 & -1 & 5 \\ 2 & -6 & 4 & -3 \\ -4 & 12 & 2 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 2 & -4 \\ 0 & 0 & 5 & -7 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & -2 & 9 & 5 & 4 \\ 1 & -1 & 6 & 5 & -3 \\ -2 & 0 & -6 & 1 & -2 \\ 4 & 1 & 9 & 1 & -9 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & -2 & 9 & 5 & 4 \\ 0 & 1 & -3 & 0 & -7 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

11. $A = \begin{bmatrix} 1 & 2 & -5 & 0 & -1 \\ 2 & 5 & -8 & 4 & 3 \\ -3 & -9 & 9 & -7 & -2 \\ 3 & 10 & -7 & 11 & 7 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & 2 & -5 & 0 & -1 \\ 0 & 1 & 2 & 4 & 5 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 2 & -4 & 3 & 3 \\ 5 & 10 & -9 & -7 & 8 \\ 4 & 8 & -9 & -2 & 7 \\ -2 & -4 & 5 & 0 & -6 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & 2 & -4 & 3 & 3 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 13 y 14, encuentre una base para el subespacio que generan los vectores dados. ¿Cuál es la dimensión del subespacio?

13. $\begin{bmatrix} 1 \\ -3 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ 9 \\ -6 \\ 12 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -3 \\ 7 \end{bmatrix}$

14. $\begin{bmatrix} 1 \\ -1 \\ -2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ -1 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -6 \\ 8 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ -7 \\ 7 \end{bmatrix}, \begin{bmatrix} 3 \\ -8 \\ 9 \\ -5 \end{bmatrix}$

15. Suponga que una matriz A de 3×5 tiene tres columnas pivote. ¿Es Col $A = \mathbb{R}^3$? ¿Es Nul $A = \mathbb{R}^2$? Explique sus respuestas.

16. Suponga que una matriz A de 4×7 tiene tres columnas pivote ¿Es Col $A = \mathbb{R}^3$? ¿Cuál es la dimensión de Nul A? Explique sus respuestas.

En los ejercicios 17 y 18, señale cada enunciado como verdadero o falso. Justifique sus respuestas. Aquí A es una matriz $m \times n$.

17. a. Si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es una base para un subespacio H, y si $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, entonces $c_1, \ldots, cp$ son las coordenadas de $\mathbf{x}$ relativas a la base $\mathcal{B}$.

b. Cada línea en $\mathbb{R}^n$ es un subespacio unidimensional de $\mathbb{R}^n$.

c. La dimensión de Col A es el número de columnas pivote de A.

d. Las dimensiones de Col A y Nul A suman el número de columnas de A.

e. Si un conjunto de p vectores genera un subespacio p-dimensional H de $\mathbb{R}^n$, entonces estos vectores forman una base para H.

18. a. Si $\mathcal{B}$ es una base para un subespacio H, entonces cada vector en H puede escribirse sólo de una forma como combinación lineal de los vectores en $\mathcal{B}$.

b. Si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es una base para un subespacio H de $\mathbb{R}^n$, entonces la correspondencia $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ hace que H se vea y actúe igual que $\mathbb{R}^p$.

c. La dimensión de Nul A es el número de variables en la ecuación $A\mathbf{x} = \mathbf{0}$.

d. La dimensión del espacio columna de A es rango A.

e. Si H es un subespacio p-dimensional de $\mathbb{R}^n$, entonces un conjunto linealmente independiente de p vectores en H es una base para H.

En los ejercicios 19 a 24 justifique cada respuesta o construcción.

19. Si el subespacio de todas las soluciones de $A\mathbf{x} = \mathbf{0}$ tiene una base que consiste en tres vectores, y si A es una matriz de 5×7, ¿cuál es el rango de A?

20. ¿Cuál es el rango de una matriz de 4×5 cuyo espacio nulo es tridimensional?

21. Si el rango de una matriz A de 7×6 es 4, ¿cuál es la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$?

22. Muestre que un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_5\}$ en $\mathbb{R}^n$ es linealmente dependiente si dim Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_5\} = 4$.

23. Si es posible, construya una matriz A de 3×4 tal que dim Nul $A = 2$ y dim Col $A = 2$.

24. Construya una matriz de 4×3 con rango 1.

25. Sea A una matriz $n \times p$ cuyo espacio columna es p-dimensional. Explique por qué las columnas de A deben ser linealmente independientes.

26. Suponga que las columnas 1, 3, 5 y 6 de una matriz A son linealmente independientes (pero no son necesariamente columnas pivote), y que el rango de A es 4. Explique por qué las cuatro columnas mencionadas deben ser una base para el espacio columna de A.

27. Suponga que los vectores $\mathbf{b}_1, \ldots, \mathbf{b}_p$ generan un subespacio W, y sea $\{\mathbf{a}_1, \ldots, \mathbf{a}_q\}$ cualquier conjunto en W que contenga más de p vectores. Complete los detalles del siguiente argumento para demostrar que $\{\mathbf{a}_1, \ldots, \mathbf{a}_q\}$ debe ser linealmente dependiente. Primero, sea $\mathcal{B} = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_p]$ y $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_q]$.

a. Explique por qué para cada vector $\mathbf{a}_j$ existe un vector $\mathbf{c}_j$ en $\mathbb{R}^p$ tal que $\mathbf{a}_j = \mathcal{B}\mathbf{c}_j$.

b. Sea $C = [\mathbf{c}_1 \ \cdots \ \mathbf{c}_q]$. Explique por qué existe un vector diferente de cero tal que $C\mathbf{u} = \mathbf{0}$.

c. Utilice $\mathcal{B}$ y C para demostrar que $A\mathbf{u} = \mathbf{0}$. Esto deja ver que las columnas de A son linealmente dependientes.

28. Use el ejercicio 27 para mostrar que si $\mathcal{A}$ y $\mathcal{B}$ son bases para un subespacio W de $\mathbb{R}^n$, entonces $\mathcal{A}$ no puede contener más vectores que $\mathcal{B}$ y, recíprocamente, que $\mathcal{B}$ no puede contener más vectores que $\mathcal{A}$.

29. **[M]** Sean $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$, y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Muestre que $\mathbf{x}$ está en H, y encuentre el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$, cuando

$$\mathbf{v}_1 = \begin{bmatrix} 11 \\ -5 \\ 10 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 14 \\ -8 \\ 13 \\ 10 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 19 \\ -13 \\ 18 \\ 15 \end{bmatrix}$$

30. **[M]** Sean $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Muestre que $\mathcal{B}$ es una base para H y que $\mathbf{x}$ está en H, también encuentre el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$, cuando

$$\mathbf{v}_1 = \begin{bmatrix} -6 \\ 4 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ -3 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 5 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 7 \\ -8 \\ 3 \end{bmatrix}$$

SG Dominio de dimensión y rango 2 a 41 (Mastering: Dimension and Rank 2-41)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Construya $A = [\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \mathbf{v}_3]$ de manera que el subespacio generado por $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ sea el espacio columna de A. Las columnas pivote de A proporcionan una base para este espacio.

$$A = \begin{bmatrix} 2 & 3 & -1 \\ -8 & -7 & 6 \\ 6 & -1 & -7 \end{bmatrix} \sim \begin{bmatrix} 2 & 3 & -1 \\ 0 & 5 & 2 \\ 0 & -10 & -4 \end{bmatrix} \sim \begin{bmatrix} 2 & 3 & -1 \\ 0 & 5 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Las primeras dos columnas de A son columnas pivote y forman una base para H. Por lo tanto, dim $H = 2$.

2. Si $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, entonces **x** se forma a partir de una combinación lineal de los vectores de la base usando los pesos 3 y 2:

$$\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2 = 3\begin{bmatrix} 1 \\ .2 \end{bmatrix} + 2\begin{bmatrix} .2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 2.6 \end{bmatrix}$$

La base $\{\mathbf{b}_1, \mathbf{b}_2\}$ determina un *sistema de coordenadas* para $\mathbb{R}^2$, lo cual se ilustra con la malla de la figura. Observe cómo **x** tiene 3 unidades en la dirección $\mathbf{b}_1$ y dos unidades en la dirección $\mathbf{b}_2$.

3. Un subespacio cuatridimensional contendría una base de cuatro vectores linealmente independientes. Esto es imposible en $\mathbb{R}^2$. Como cualquier conjunto linealmente independiente en $\mathbb{R}^3$ no contiene más de tres vectores, cualquier subespacio de $\mathbb{R}^3$ tiene una dimensión no mayor a 3. El propio espacio $\mathbb{R}^3$ es el único subespacio tridimensional de $\mathbb{R}^3$. Los otros subespacios de $\mathbb{R}^3$ tienen dimensión 2, 1 o 0.

CAPÍTULO 2 EJERCICIOS SUPLEMENTARIOS

1. Suponga que las matrices mencionadas en los siguientes enunciados tienen los tamaños adecuados. Señale cada enunciado como verdadero o falso. Justifique sus respuestas
   a. Si A y B son $m \times n$, entonces tanto AB^T como A^TB están definidas.
   b. Si $AB = C$ y C tiene dos columnas, entonces A tiene dos columnas.
   c. Al multiplicar por la izquierda una matriz B por una matriz diagonal A, con entradas distintas de cero en la diagonal, se escalan las filas de B.
   d. Si $BC = BD$, entonces $C = D$.
   e. Si $AC = 0$, entonces $A = 0$ o bien $C = 0$.
   f. Si A y B son $n \times n$, entonces $(A + B)(A - B) = A^2 - B^2$.
   g. Una matriz elemental de $n \times n$ tiene o n o bien $n + 1$ entradas diferentes de cero.
   h. La transpuesta de una matriz elemental es una matriz elemental.
   i. Una matriz elemental debe ser cuadrada.
   j. Toda matriz cuadrada es un producto de matrices elementales.
   k. Si A es una matriz de 3×3 con tres posiciones pivote, existen matrices elementales $E_1, \ldots, E_p$ tales que $E_p \cdots E_1A = I$.
   l. Si $AB = I$, entonces A es invertible.
   m. Si A y B son cuadradas e invertibles, entonces AB es invertible, y $(AB)^{-1} = A^{-1}B^{-1}$.
   n. Si $AB = BA$ y A es invertible, entonces $A^{-1}B = BA^{-1}$.
   o. Si A es invertible y $r \neq 0$, entonces $(rA)^{-1} = rA^{-1}$.
   p. Si A es una matriz de 3×3 y la ecuación $A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ tiene una sola solución, entonces A es invertible.

2. Encuentre la matriz C cuyo inverso es $C^{-1} = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$.

3. Sea $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Muestre que $A^3 = 0$. Utilice álgebra matricial para calcular el producto $(I - A)(I + A + A^2)$.

4. Suponga que $A^n = 0$ para alguna $n > 1$. Encuentre un inverso de $I - A$.

5. Suponga que una matriz A de $n \times n$ satisface la ecuación $A^2 - 2A + I = 0$. Muestre que $A^3 = 3A - 2I$, y que $A^4 = 4A - 3I$.

6. Sea $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Éstas son *matrices de espín de Pauli* y se usan en mecánica cuántica para el estudio del espín de electrones. Demuestre que $A^2 = I$, $B^2 = I$ y $AB = -BA$. Matrices del tipo $AB = -BA$ se llaman *anticonmutativas.*

7. Sea $A = \begin{bmatrix} 1 & 3 & 8 \\ 2 & 4 & 11 \\ 1 & 2 & 5 \end{bmatrix}$ y $B = \begin{bmatrix} -3 & 5 \\ 1 & 5 \\ 3 & 4 \end{bmatrix}$. Determine $A^{-1}B$ sin calcular A^{-1}. [*Indicación:* $A^{-1}B$ es la solución de la ecuación $AX = B$.]

8. Encuentre una matriz A tal que la transformación $\mathbf{x} \to A\mathbf{x}$ mapee $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 2 \\ 7 \end{bmatrix}$ en $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, respectivamente. [*Sugerencia:* Escriba una ecuación de matrices que contenga a A, y despeje A.]

9. Suponga que $AB = \begin{bmatrix} 5 & 4 \\ -2 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix}$. Encuentre A.

10. Suponga que A es invertible. Explique por qué A^TA también es invertible. Luego demuestre que $A^{-1} = (A^TA)^{-1}A^T$.

11. Sean $x_1, \ldots, x_n$ números fijos. La matriz siguiente, llamada *matriz de Vandermonde*, aparece en aplicaciones como procesamiento de señales, códigos correctores de errores, e interpolación de polinomios.

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

Dado $\mathbf{y} = (y_1, \ldots, y_n)$ en $\mathbb{R}^n$, suponga que $\mathbf{c} = (c_0, \ldots, c_{n-1})$ en $\mathbb{R}^n$ satisface $V\mathbf{c} = \mathbf{y}$, y defina el polinomio

$$p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_{n-1}t^{n-1}$$

a. Demuestre que $p(x_1) = y_1, \ldots, p(x_n) = y_n$. Se llama a $p(t)$ un *polinomio de interpolación para los puntos* $(x_1, y_1), \ldots, (x_n, y_n)$ porque la gráfica de $p(t)$ pasa por estos puntos.

b. Suponga que $x_1, \ldots, x_n$ son números distintos. Muestre que las columnas de V son linealmente independientes. [*Sugerencia:* ¿Cuántos ceros puede tener un polinomio de grado $n - 1$?]

c. Demuestre que: "Si $x_1, \ldots, x_n$ son números distintos y $y_1, \ldots, y_n$ son números arbitrarios, entonces hay un polinomio de interpolación de grado $\leq n - 1$ para $(x_1, y_1), \ldots, (x_n, y_n)$."

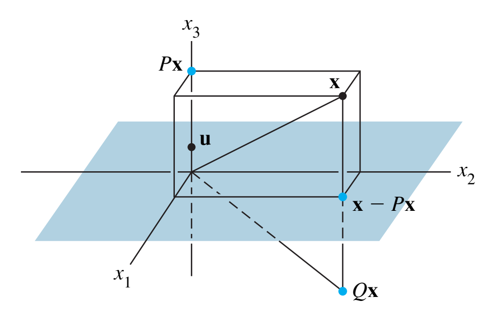
12. Sea $A = LU$, donde L es una matriz triangular inferior invertible y U es triangular superior. Explique por qué la primera columna de A es un múltiplo de la primera columna de L. ¿Cómo se relaciona la segunda columna de A con las columnas de L?

13. Dado $\mathbf{u}$ en $\mathbb{R}^n$ con $\mathbf{u}^T\mathbf{u} = 1$, sea $P = \mathbf{u}\mathbf{u}^T$ (un producto exterior) y $Q = I - 2P$. Justifique los enunciados (a), (b) y (c).

a. $P^2 = P$ b. $P^T = P$ c. $Q^2 = I$

La transformación $\mathbf{x} \mapsto P\mathbf{x}$ es una *proyección*, y $\mathbf{x} \mapsto Q\mathbf{x}$ se llama *reflexión de Householder*. Tales reflexiones se usan en programas de computadora para crear múltiples ceros en un vector (por lo general, una columna de una matriz).

14. Sea $\mathbf{u} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$. Determine P y Q igual que en el ejercicio 13, y calcule $P\mathbf{x}$ y $Q\mathbf{x}$. La figura muestra que $Q\mathbf{x}$ es la reflexión de $\mathbf{x}$ a través del plano x_1x_2.



Una reflexión de Householder a través del plano $x_3 = 0$.

15. Suponga que $C = E_3E_2E_1B$, donde E_1, E_2, E_3 son matrices elementales. Explique por qué C es equivalente por filas a B.

16. Sea A una matriz singular de $n \times n$. Describa cómo puede construirse una matriz B, de $n \times n$, diferente de cero tal que $AB = 0$.

17. Sean A una matriz de 6×4 y B una matriz de 4×6. Muestre que la matriz AB de 6×6 no puede ser invertible.

18. Suponga que A es una matriz de 5×3 y que existe una matriz C de 3×5 tal que $CA = I_3$. Suponga además que para alguna $\mathbf{b}$ dada en $\mathbb{R}^5$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene por lo menos una solución. Muestre que esta solución es única.

19. **[M]** Ciertos sistemas dinámicos se pueden estudiar examinando las potencias de una matriz, como las presentadas a continuación. Determine qué les pasa a A^k y B^k conforme se incrementa k (por ejemplo, pruebe con $k = 2, \ldots, 16$). Trate de identificar qué tienen de especial A y B. Investigue potencias grandes de otras matrices de este tipo y formule una conjetura acerca de tales matrices.

$$A = \begin{bmatrix} .4 & .2 & .3 \\ .3 & .6 & .3 \\ .3 & .2 & .4 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & .2 & .3 \\ .1 & .6 & .3 \\ .9 & .2 & .4 \end{bmatrix}$$

20. **[M]** Sea A_n una matriz $n \times n$ con ceros en la diagonal principal y números 1 en el resto. Calcule A_n^{-1} para $n = 4, 5$ y 6, y formule una conjetura acerca de la forma general de A_n^{-1} para valores más grandes de n.

CD Interpolación de Lagrange (Lagrange Interpolation)

3

Determinantes

EJEMPLO INTRODUCTORIO

Determinantes en geometría analítica

Un determinante es un número que se asigna de cierto modo a una formación cuadrada de números. Esta idea fue considerada en 1683 por el matemático japonés Seki Takakasu y, de manera independiente, en 1693 por el matemático alemán Gottfried Leibniz, unos 160 años antes de que se desarrollara una teoría de matrices por separado. Durante muchos años, los determinantes aparecieron principalmente en relación con sistemas de ecuaciones lineales.

En 1750, un artículo del matemático suizo Gabriel Cramer sugirió que los determinantes podrían ser útiles en geometría analítica. En ese documento, Cramer usó determinantes para construir ecuaciones de ciertas curvas en el plano xy. En el mismo texto, también presentó su famosa regla para resolver un sistema $n \times n$ mediante determinantes. Después, en 1812, Augustin-Louis Cauchy publicó un documento donde utilizó determinantes con el propósito de encontrar fórmulas para los volúmenes de ciertos poliedros sólidos, y estableció una conexión entre dichas fórmulas y los trabajos previos sobre determinantes. Entre los "cristales" que estudió Cauchy estaban el tetraedro de la figura 1 y el paralelepípedo de la figura 2. Si los vértices del paralelepípedo son el origen $\mathbf{0} = (0, 0, 0)$, $\mathbf{v}_1 = (a_1, b_1, c_1)$, $\mathbf{v}_2 = (a_2, b_2, c_2)$, y $\mathbf{v}_3 = (a_3, b_3, c_3)$, entonces su volumen es el valor absoluto del determinante de la matriz de coeficientes del sistema:

$$\begin{aligned} a_1x + b_1y + c_1z &= 0 \\ a_2x + b_2y + c_2z &= 0 \\ a_3x + b_3y + c_3z &= 0 \end{aligned}$$

El uso que en geometría analítica hizo Cauchy de los determinantes despertó un profundo interés en las aplicaciones de los determinantes, lo cual duró aproximadamente 100 años. Un simple resumen de lo que se conocía a principios del siglo XX llenó un tratado de cuatro volúmenes escrito por Thomas Muir.

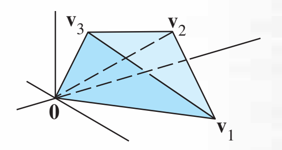


FIGURA 1 Un tetraedro.

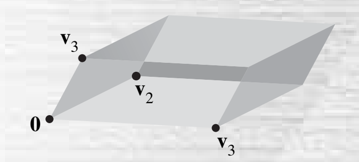


FIGURA 2 Un paralelepípedo.

En tiempos de Cauchy, cuando la vida era simple y las matrices eran pequeñas, los determinantes desempeñaron un papel importante en geometría analítica y en otras áreas de las matemáticas. En la actualidad, los determinantes tienen escaso valor numérico en los cálculos de matrices a gran escala que surgen con frecuencia. No obstante, las fórmulas para determinantes todavía proporcionan información importante acerca de las matrices y el conocimiento de los determinantes resulta útil en algunas aplicaciones del álgebra lineal.

Las metas para este capítulo son tres: demostrar un criterio de invertibilidad para una matriz cuadrada A en el que intervienen las entradas de A en vez de sus columnas; proporcionar fórmulas para A^{-1} y $A^{-1}\mathbf{b}$ que se usan en aplicaciones teóricas, y deducir la interpretación geométrica de los determinantes descritos en la introducción del capítulo. En la sección 3.2 se alcanza la primera meta, y en la sección 3.3 las dos restantes.

3.1 INTRODUCCIÓN A LOS DETERMINANTES

De la sección 2.2, recuerde que una matriz de 2×2 es invertible si, y sólo si, su determinante es diferente de cero. Para extender este útil hecho a matrices más grandes, se requiere una definición para el determinante de una matriz $n \times n$. Es posible descubrir la definición para el caso 3×3 al observar lo que sucede cuando se reduce por filas una matriz invertible A de 3×3.

Considere $A = [a_{ij}]$ con $a_{11} \neq 0$. Si se multiplican la segunda y tercera filas de A por a_{11} y luego se restan múltiplos apropiados de la primera fila a las otras dos filas, se encuentra que A es equivalente por filas a las siguientes dos matrices:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11}a_{21} & a_{11}a_{22} & a_{11}a_{23} \\ a_{11}a_{31} & a_{11}a_{32} & a_{11}a_{33} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & a_{11}a_{32} - a_{12}a_{31} & a_{11}a_{33} - a_{13}a_{31} \end{bmatrix} \qquad (1)$$

Como A es invertible, la entrada (2, 2) o bien la entrada (3, 2) a la derecha de (1) es diferente de cero. Suponga que la entrada (2, 2) es diferente de cero. (De lo contrario, puede hacerse un intercambio de filas antes de proseguir.) Multiplique la fila 3 por $a_{11}a_{22} - a_{12}a_{21}$, y luego sume a la nueva fila 3 $-(a_{11}a_{32} - a_{12}a_{31})$ veces la fila 2. Esto demostrará que

$$A \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & 0 & a_{11}\Delta \end{bmatrix}$$

donde

$$\Delta = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \qquad (2)$$

Como A es invertible, Δ debe ser diferente de cero. El recíproco también es cierto, como se verá en la sección 3.2. El valor Δ en (2) se llama **determinante** de la matriz A de 3×3.

Recuerde que el determinante de una matriz de 2×2, $A = [a_{ij}]$, es el número

$$\det A = a_{11}a_{22} - a_{12}a_{21}$$

Para una matriz de 1×1 —por ejemplo, $A = [a_{11}]$— se define $\det A = a_{11}$. Para generalizar la definición del determinante para matrices más grandes, se utilizarán determinantes de 2×2 para reescribir el determinante Δ 3×3 descrito con anterioridad. Dado que los términos de Δ pueden agruparse como

$$(a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32}) - (a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31}) + (a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}),$$

$$\Delta = a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

Por brevedad, se escribe

$$\Delta = a_{11} \cdot \det A_{11} - a_{12} \cdot \det A_{12} + a_{13} \cdot \det A_{13} \tag{3}$$

donde A_{11}, A_{12} y A_{13} se obtienen de A al eliminar la primera fila y una de las tres columnas. Para cualquier matriz cuadrada A, A_{ij} denotará la submatriz formada al borrar la i-ésima fila y la j-ésima columna de A. Por ejemplo, si

$$A = \begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix}$$

entonces A_{32} se obtiene tachando la fila 3 y la columna 2,

$$\begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix}$$

de manera que

$$A_{32} = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Ahora puede darse una definición *recursiva* de un determinante. Cuando $n = 3$, det A se define usando determinantes de las submatrices A_{1j} de 2×2, como ya se vio en (3). Cuando $n = 4$, det A utiliza los determinantes de las submatrices A_{1j} de 3×3. En general, un determinante $n \times n$ se define mediante determinantes de submatrices $(n - 1) \times (n - 1)$.

DEFINICIÓN

Para $n \geq 2$, el **determinante** de una matriz A de $n \times n = [a_{ij}]$ es la suma de los n términos de la forma $\pm a_{1j} \det A_{1j}$, con los signos más y menos alternándose, donde las entradas $a_{11}, a_{12}, ..., a_{1n}$ son de la primera fila de A. En forma simbólica,

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n}$$
$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$$

EJEMPLO 1 Calcule el determinante de

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Solución Calcule $\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}$:

$$\begin{aligned}\det A &= 1 \cdot \det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5 \cdot \det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0 \cdot \det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix} \\ &= 1(0-2) - 5(0-0) + 0(-4-0) = -2\end{aligned}$$

Otra notación común para el determinante de una matriz usa un par de líneas verticales en lugar de los corchetes. Así, el cálculo del ejemplo 1 se puede escribir como

$$\det A = 1 \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5 \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0 \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} = \cdots = -2$$

Para enunciar el teorema siguiente resulta oportuno escribir la definición $\det A$ en una forma un poco diferente. Dada $A = [a_{ij}]$, el **cofactor (i, j)** de A es el número C_{ij} dado por

$$C_{ij} = (-1)^{i+j} \det A_{ij} \tag{4}$$

Entonces

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$

Esta fórmula se llama **desarrollo por cofactores a lo largo de la primera fila** de A. Se omite la demostración del teorema fundamental siguiente para evitar una larga interrupción.

TEOREMA 1

El determinante de una matriz A de $n \times n$ puede calcularse mediante un desarrollo por cofactores a lo largo de cualquier fila o descendiendo por cualquier columna. El desarrollo a lo largo de la i-ésima fila usando los cofactores en (4) es

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

El desarrollo por cofactores bajando por la j-ésima columna es

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

Los signos más o menos del cofactor (i, j) dependen de la posición de a_{ij} en la matriz, sin importar el signo de a_{ij} en sí mismo. El factor $(-1)^{i+j}$ determina la tabla siguiente para el patrón de signos:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

EJEMPLO 2 Use un desarrollo por cofactores a lo largo de la tercera fila para calcular $\det A$, donde

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Solución Calcule

$$\begin{aligned}\det A &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \\ &= (-1)^{3+1}a_{31}\det A_{31} + (-1)^{3+2}a_{32}\det A_{32} + (-1)^{3+3}a_{33}\det A_{33} \\ &= 0\begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} \\ &= 0 + 2(-1) + 0 = -2\end{aligned}$$

El teorema 1 es útil para calcular los determinantes de una matriz que contiene muchos ceros. Por ejemplo, si una fila está formada en su mayoría por ceros, entonces el desarrollo por cofactores a lo largo de esa fila tiene muchos términos que son cero, y no es necesario calcular los cofactores en esos términos. El mismo enfoque funciona con una columna que contiene muchos ceros.

EJEMPLO 3 Calcule det A, donde

$$A = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$$

Solución El desarrollo por cofactores bajando por la primera columna de A tiene todos los términos iguales a cero excepto el primero. Entonces

$$\det A = 3\cdot\begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} - 0\cdot C_{21} + 0\cdot C_{31} - 0\cdot C_{41} + 0\cdot C_{51}$$

A partir de aquí se omitirán los términos cero en el desarrollo por cofactores. Enseguida, desarrolle este determinante 4×4 descendiendo por la primera columna, para aprovechar los ceros que existen ahí. Se tiene

$$\det A = 3\cdot 2\cdot\begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}$$

Este determinante 3×3 se calculó en el ejemplo 1 y se vio que era igual a -2. Así que $\det A = 3 \cdot 2 \cdot (-2) = -12$.

La matriz del ejemplo 3 era casi triangular. El método de ese ejemplo puede adaptarse con facilidad para demostrar el teorema siguiente.

TEOREMA 2 Si A es una matriz triangular, entonces det A es el producto de las entradas sobre la diagonal principal de A.

La estrategia del ejemplo 3 de buscar ceros funciona extremadamente bien cuando toda una fila o columna consiste en ceros. En un caso así, el desarrollo por cofactores a lo largo de una fila o columna de este tipo es una suma de ceros. Así que el determinante es cero. Desafortunadamente, la mayor parte de los desarrollos por cofactores no se calcula con tanta rapidez.

NOTA NUMÉRICA

Para los estándares actuales, una matriz de 25×25 es pequeña. Aún así, resultaría imposible calcular un determinante 25×25 empleando el desarrollo por cofactores. En general, un desarrollo por cofactores requiere más de $n!$ multiplicaciones, y $25!$ es aproximadamente 1.5×10^{25}.

Si una computadora realizara un trillón de multiplicaciones por segundo, tendría que trabajar durante más de 500,000 años para calcular un determinante 25×25 con este método. Afortunadamente, como se descubrirá en breve, hay métodos más rápidos.

En los ejercicios 19 a 38 se exploran importantes propiedades de los determinantes, en su mayor parte para el caso de 2×2. Los resultados de los ejercicios 33 a 36 se utilizarán en la siguiente sección para derivar las propiedades análogas de matrices $n \times n$.

PROBLEMA DE PRÁCTICA

Calcule $\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix}$.

3.1 EJERCICIOS

En los ejercicios 1 a 8, calcule el determinante utilizando un desarrollo por cofactores a lo largo de la primera fila. En los ejercicios 1 a 4, calcule también el determinante aplicando un desarrollo por cofactores y bajando por la segunda columna.

1. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$ **2.** $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

3. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$ **4.** $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

5. $\begin{vmatrix} 2 & 3 & -4 \\ 4 & 0 & 5 \\ 5 & 1 & 6 \end{vmatrix}$ **6.** $\begin{vmatrix} 5 & -2 & 4 \\ 0 & 3 & -5 \\ 2 & -4 & 7 \end{vmatrix}$

7. $\begin{vmatrix} 4 & 3 & 0 \\ 6 & 5 & 2 \\ 9 & 7 & 3 \end{vmatrix}$ **8.** $\begin{vmatrix} 8 & 1 & 6 \\ 4 & 0 & 3 \\ 3 & -2 & 5 \end{vmatrix}$

Encuentre los determinantes en los ejercicios 9 a 14 mediante desarrollo de cofactores. En cada paso, elija una fila o columna que implique la menor cantidad de cálculos.

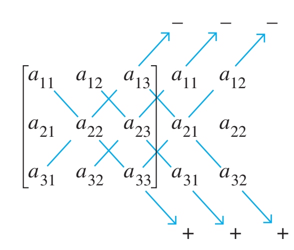
9. $\begin{vmatrix} 6 & 0 & 0 & 5 \\ 1 & 7 & 2 & -5 \\ 2 & 0 & 0 & 0 \\ 8 & 3 & 1 & 8 \end{vmatrix}$ **10.** $\begin{vmatrix} 1 & -2 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 2 & -6 & -7 & 5 \\ 5 & 0 & 4 & 4 \end{vmatrix}$

11. $\begin{vmatrix} 3 & 5 & -8 & 4 \\ 0 & -2 & 3 & -7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$ **12.** $\begin{vmatrix} 4 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ 2 & 6 & 3 & 0 \\ 5 & -8 & 4 & -3 \end{vmatrix}$

13. $\begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 0 & 2 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix}$

14. $\begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$

El desarrollo de un determinante 3 × 3 puede recordarse usando el siguiente esquema. Escriba una segunda copia de las primeras dos columnas ubicadas a la derecha de la matriz, y calcule el determinante multiplicando las entradas de seis diagonales:



Sume los productos diagonales descendentes y reste los productos ascendentes. Use este método para calcular los determinantes de los ejercicios 15 a 18. ***Advertencia:*** *Este truco no se generaliza de ninguna manera razonable a matrices de* 4 × 4 *o mayores.*

15. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$

16. $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

17. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$

18. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

En los ejercicios 19 a 24, indague el efecto de una operación elemental de fila sobre el determinante de una matriz. En cada caso, enuncie la operación de fila y describa cómo afecta al determinante.

19. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} c & d \\ a & b \end{bmatrix}$

20. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a & b \\ kc & kd \end{bmatrix}$

21. $\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 4 \\ 5+3k & 6+4k \end{bmatrix}$

22. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a+kc & b+kd \\ c & d \end{bmatrix}$

23. $\begin{bmatrix} 1 & 1 & 1 \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}, \begin{bmatrix} k & k & k \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}$

24. $\begin{bmatrix} a & b & c \\ 3 & 2 & 2 \\ 6 & 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 2 & 2 \\ a & b & c \\ 6 & 5 & 6 \end{bmatrix}$

Encuentre los determinantes de las matrices elementales dadas en los ejercicios 25 a 30. (Vea la sección 2.2.)

25. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$

26. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}$

27. $\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

28. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$

29. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

30. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

Use los ejercicios 25 a 28 para contestar las preguntas 31 y 32 siguientes. Proporcione las razones de sus respuestas.

31. ¿Cuál es el determinante de una matriz elemental de reemplazo por fila?

32. ¿Cuál es el determinante de una matriz elemental escalonada con k en la diagonal?

En los ejercicios 33 a 36, verifique que det EA = (det E)(det A), donde E es la matriz elemental que se muestra y $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

33. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

34. $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

35. $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

36. $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

37. Sea $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$. Escriba $5A$. ¿Es det $5A$ = 5 det A?

38. Sean $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ y k un escalar. Encuentre una fórmula que relacione det kA con k y det A.

En los ejercicios 39 y 40, A es una matriz $n \times n$. Señale cada afirmación como verdadera o falsa. Justifique sus respuestas.

39. a. Un determinante $n \times n$ está definido por determinantes de submatrices de $(n - 1) \times (n - 1)$.

b. El cofactor (i,j) de una matriz A es la matriz A_{ij} que se obtiene al eliminar de A su i-ésima fila y su j-ésima columna.

40. a. El desarrollo por cofactores de det A bajando por una columna es el negativo del desarrollo por cofactores a lo largo de una fila.

b. El determinante de una matriz triangular es la suma de las entradas sobre la diagonal principal.

41. Sean $\mathbf{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Calcule el área del paralelogramo determinado mediante $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u} + \mathbf{v}$, y $\mathbf{0}$, y encuentre el determinante de $[\mathbf{u} \quad \mathbf{v}]$. ¿Qué diferencias hay entre ellos?

Reemplace la primera entrada de **v** por un número arbitrario x, y repita el problema. Trace un dibujo y explique lo que encuentre.

42. Sean $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} c \\ 0 \end{bmatrix}$, donde a, b y c son positivos (para simplificar). Calcule el área del paralelogramo determinado por **u**, **v**, **u** + **v**, y **0**, y encuentre los determinantes de las matrices [**u** **v**] y [**v** **u**]. Trace un dibujo y explique lo que encuentre.

43. **[M]** ¿Es cierto que $\det(A + B) = \det A + \det B$? Para responder esta pregunta, genere matrices aleatorias de 5×5 A y B, y calcule $\det(A + B) - \det A - \det B$. (Remítase al ejercicio 37 de la sección 2.1.) Repita los cálculos para otros pares de matrices $n \times n$, con diferentes valores de n. Informe acerca de sus resultados.

44. **[M]** ¿Es cierto que $\det AB = (\det A)(\det B)$? Experimente con cuatro pares de matrices aleatorias, como en el ejercicio 43, y formule una conjetura.

45. **[M]** Construya una matriz aleatoria A de 4×4 con entradas enteras entre -9 y 9, y compare $\det A$ con $\det A^T$, $\det(-A)$, $\det(2A)$, y $\det(10A)$. Repita con otras dos matrices aleatorias de 4×4, y formule conjeturas acerca de la relación entre estos determinantes. (Remítase al ejercicio 36 de la sección 2.1.) Luego compruebe sus conjeturas con varias matrices aleatorias enteras de 5×5 y 6×6. Si es necesario, modifique sus conjeturas e informe los resultados.

46. **[M]** ¿Qué relación hay entre $\det A^{-1}$ y $\det A$? Experimente con matrices aleatorias enteras *de* $n \times n$, con $n = 4$, 5 y 6, y formule una conjetura. *Nota:* En el caso poco probable de que encuentre una matriz con determinante cero, redúzcala a su forma escalonada y analice lo que encuentre.

Solución al problema de práctica

Aproveche los ceros. Empiece con un desarrollo por cofactores descendiendo por la tercera columna para obtener una matriz de 3×3, la cual puede calcularse por medio de un desarrollo descendiendo por su primera columna.

$$\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix} = (-1)^{1+3}2 \begin{vmatrix} 0 & 3 & -4 \\ -5 & -8 & 3 \\ 0 & 5 & -6 \end{vmatrix}$$

$$= 2\cdot(-1)^{2+1}(-5) \begin{vmatrix} 3 & -4 \\ 5 & -6 \end{vmatrix} = 20$$

El $(-1)^{2+1}$ del penúltimo cálculo proviene de la posición (2,1) del -5 localizado en el determinante de 3×3.

3.2 PROPIEDADES DE LOS DETERMINANTES

El secreto de los determinantes radica en cómo cambian cuando se realizan operaciones por fila. El teorema siguiente generaliza los resultados de los ejercicios 19 a 24 de la sección 3.1. La demostración aparece al final de este apartado.

TEOREMA 3 **Operaciones por fila**

Sea A una matriz cuadrada.

a. Si un múltiplo de una fila de A se suma a otra fila para producir una matriz B, entonces $\det B = \det A$.

b. Si dos filas de A se intercambian para producir B, entonces $\det B = -\det A$.

c. Si una fila de A se multiplica por k para producir B, entonces $\det B = k \cdot \det A$.

En los siguientes ejemplos se muestra cómo usar el teorema 3 para encontrar los determinantes de manera eficiente.

EJEMPLO 1 Calcule det A, donde $A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.

Solución La estrategia es reducir A a la forma escalonada y utilizar luego el hecho de que el determinante de una matriz triangular es el producto de las entradas diagonales. Los primeros dos reemplazos de fila en la columna 1 no alteran el determinante:

$$\det A = \begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix}$$

Un intercambio de las filas 2 y 3 invierte el signo del determinante, así que

$$\det A = -\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} = -(1)(3)(-5) = 15$$

En cálculos hechos a mano, un uso común del teorema 3(c) es el de *encontrar un factor que sea un múltiplo común de una fila* de una matriz. Por ejemplo,

$$\begin{vmatrix} * & * & * \\ 5k & -2k & 3k \\ * & * & * \end{vmatrix} = k \begin{vmatrix} * & * & * \\ 5 & -2 & 3 \\ * & * & * \end{vmatrix}$$

donde las entradas con asterisco no cambian. Este paso se utilizará en el siguiente ejemplo.

EJEMPLO 2 Calcule det A, donde A $A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$.

Solución Para simplificar la aritmética, se desea un 1 en la esquina superior izquierda. Se podrían intercambiar las filas 1 y 4. En lugar de eso, se saca el factor 2 de la fila superior y se procede con los reemplazos de fila en la primera columna:

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Luego, se podría sacar otro factor 2 de la fila 3 o usar el 3 de la segunda columna como un pivote. Se elige la última operación, sumando 4 veces la fila 2 a la fila 3:

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Por último, al sumar $-1/2$ veces la fila 3 a la fila 4, y al calcular el determinante "triangular", se tiene que

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2\cdot(1)(3)(-6)(1) = -36$$

$$U = \begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{bmatrix}$$
$\det U \neq 0$

$$U = \begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$\det U = 0$

FIGURA 1
Formas escalonadas típicas de matrices cuadradas.

Suponga que una matriz cuadrada A se ha reducido a una forma escalonada U mediante el reemplazo e intercambio de filas. (Esto siempre es posible. Vea el algoritmo de reducción por filas de la sección 1.2.) Si hay r intercambios, entonces el teorema 3 muestra que

$$\det A = (-1)^r \det U$$

Como U está en forma escalonada, es triangular, y también $\det U$ es el producto de las entradas diagonales $u_{11}, \ldots, u_{nn}$. Si A es invertible, las entradas u_{ii} son todas pivotes (porque $A \sim I_n$ y las u_{ii} no se han escalado a números 1). De lo contrario, al menos u_{nn} es cero, y el producto $u_{11} \cdots u_{nn}$ es cero. Vea la figura 1. Así que

$$\det A = \begin{cases} (-1)^r \cdot \begin{pmatrix} \text{producto de los} \\ \text{pivotes en } U \end{pmatrix} & \text{cuando } A \text{ es invertible} \\ 0 & \text{cuando } A \text{ no es invertible} \end{cases} \quad (1)$$

Resulta interesante advertir que, aunque la forma escalonada no es única (porque no está completamente reducida por filas) y los pivotes no son únicos, el *producto* de los pivotes *sí es* único, excepto por un posible signo menos.

La fórmula (1) proporciona no sólo una interpretación muy concreta de lo que es un determinante, sino que también demuestra el teorema principal de esta sección:

TEOREMA 4 Una matriz cuadrada A es invertible si, y sólo si, $\det A \neq 0$.

El teorema 4 añade al teorema de la matriz invertible el enunciado "$\det A \neq 0$". Un corolario útil es que $\det A = 0$ cuando las columnas de A son linealmente dependientes. También, $\det A = 0$ cuando las *filas* de A son linealmente dependientes. (Las filas de A son columnas de A^T, y las columnas linealmente dependientes de A^T vuelven singular a A^T. Cuando A^T es singular, también A lo es, por el teorema de la matriz invertible.) En la práctica, la dependencia lineal sólo resulta evidente cuando dos filas o dos columnas son iguales, o cuando una fila o columna es cero.

EJEMPLO 3 Calcule $\det A$, donde $A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$.

Solución Sume 2 veces la fila 1 a la fila 3 para obtener

$$\det A = \det \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix} = 0$$

porque la segunda y tercera filas de la segunda matriz son iguales. ■

Notas numéricas

1. La mayor parte de los programas de computadora que calculan det A para una matriz general A usan el anterior método de la fórmula (1).
2. Es posible demostrar que calcular un determinante $n \times n$ usando operaciones por fila requiere de aproximadamente $2n^3/3$ operaciones aritméticas. Cualquier microcomputadora moderna puede calcular un determinante 25×25 en una fracción de segundo, puesto que sólo se requieren unas 10,000 operaciones.

CD Determinantes y flops (Determinants and Flops)

Las computadoras también pueden manejar grandes matrices "ralas" con rutinas especiales que aprovechan la presencia de muchos ceros. Por supuesto, las entradas 0 también pueden acelerar los cálculos a mano. Los cálculos del siguiente ejemplo combinan el poder de las operaciones por fila con la estrategia de la sección 3.1 de usar entradas 0 en los desarrollos por cofactores.

EJEMPLO 4 Calcule det A, donde $A = \begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{bmatrix}$.

Solución Una buena manera de comenzar es usando el 2 de la columna 1 como un pivote, eliminando el -2 que está debajo. Luego se usa un desarrollo por cofactores para reducir el tamaño del determinante, seguido de otra operación de reemplazo. Así que,

$$\det A = \begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & -3 & 1 \end{vmatrix}$$

Ahora se podrían intercambiar las filas 2 y 3 para obtener un determinante "triangular". Otro enfoque consiste en realizar un desarrollo por cofactores descendiendo por la primera columna:

$$\det A = (-2)(1) \begin{vmatrix} 0 & 5 \\ -3 & 1 \end{vmatrix} = -2 \cdot (15) = -30$$ ■

Operaciones de columna

Pueden realizarse operaciones con las columnas de una matriz de manera análoga a las operaciones por fila que se han considerado hasta el momento. El teorema siguiente muestra que las operaciones por columna tienen los mismos efectos sobre los determinantes que las operaciones por fila.

TEOREMA 5

Si A es una matriz $n \times n$, entonces $\det A^T = \det A$.

DEMOSTRACIÓN El teorema resulta evidente para $n = 1$. Suponga que el teorema es verdadero para determinantes $k \times k$ y sea $n = k + 1$. Entonces el cofactor de a_{1j} en A es igual al cofactor de a_{j1} en A^T, porque los cofactores implican determinantes $k \times k$. Por lo tanto, el desarrollo por cofactores de $\det A$ a lo largo de la primera *fila* es igual al desarrollo por cofactores de $\det A^T$ descendiendo por la primera *columna.* Es decir, A y A^T tienen determinantes iguales. Así que el teorema es cierto para $n = 1$, y su validez para un valor de n implica su validez para el siguiente valor de n. Por el principio de inducción, el teorema es cierto para toda $n \geq 1$. ■

De acuerdo con el teorema 5, cada enunciado del teorema 3 es cierto cuando la palabra *fila* se reemplaza en todas partes por la palabra *columna.* Para verificar esta propiedad, simplemente se aplica el teorema 3 original a A^T. Una operación por filas sobre A^T equivale a una operación por columnas sobre A.

Las operaciones por columna resultan útiles tanto para propósitos teóricos como para realizar cálculos a mano. Sin embargo, por simplicidad, sólo se realizarán operaciones por filas en los cálculos numéricos.

Determinantes y productos de matrices

La demostración del útil teorema que se presenta enseguida está al final de la sección. Las aplicaciones se encuentran en los ejercicios.

TEOREMA 6

Propiedad multiplicativa

Si A y B son matrices $n \times n$, entonces $\det AB = (\det A)(\det B)$.

EJEMPLO 5 Verifique el teorema 6 para $A = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$.

Solución

$$AB = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 25 & 20 \\ 14 & 13 \end{bmatrix}$$

y

$$\det AB = 25 \cdot 13 - 20 \cdot 14 = 325 - 280 = 45$$

Como $\det A = 9$ y $\det B = 5$,

$$(\det A)(\det B) = 9 \cdot 5 = 45 = \det AB$$ ■

Advertencia: Un error común es creer que el teorema 6 tiene un análogo para las *sumas* de matrices. Sin embargo, en general, $\det(A + B)$ *no* es igual a $\det A + \det B$.

Una propiedad de linealidad de la función determinante

Para una matriz A de $n \times n$, se puede considerar a $\det A$ como una función de los n vectores columna de A. Se mostrará que si todas las columnas excepto una se mantienen fijas, entonces $\det A$ es una *función lineal* de esa variable (vectorial) en particular.

Suponga que la j-ésima columna de A puede variar, y escriba

$$A = [\, \mathbf{a}_1 \quad \cdots \quad \mathbf{a}_{j-1} \quad \mathbf{x} \quad \mathbf{a}_{j+1} \quad \cdots \quad \mathbf{a}_n \,]$$

Defina una transformación T de $\mathbb{R}^n$ $\mathbb{R}$ mediante

$$T(\mathbf{x}) = \det [\, \mathbf{a}_1 \quad \cdots \quad \mathbf{a}_{j-1} \quad \mathbf{x} \quad \mathbf{a}_{j+1} \quad \cdots \quad \mathbf{a}_n \,]$$

Entonces,

$$T(c\mathbf{x}) = cT(\mathbf{x}) \quad \text{para todo escalar } c \text{ y todo } \mathbf{x} \text{ en } \mathbb{R}^n \tag{2}$$

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \quad \text{para todos } \mathbf{u}, \mathbf{v} \text{ en } \mathbb{R}^n \tag{3}$$

La propiedad (2) es el teorema 3(c) aplicado a las columnas de A. Una demostración de la propiedad (3) se sigue de un desarrollo por cofactores de $\det A$ bajando por la j-ésima columna. (Vea el ejercicio 43.) Esta propiedad de (multi-)linealidad de los determinantes resulta tener muchas consecuencias útiles, las cuales se estudian en cursos más avanzados.

Demostraciones de los teoremas 3 y 6

Es conveniente demostrar el teorema 3 cuando se plantea en términos de las matrices elementales que se estudiaron en la sección 2.2. Una matriz elemental E se denomina (*matriz de*) *reemplazo de fila* si E se obtiene a partir de la identidad I al sumar un múltiplo de una fila a otra fila; E es *de intercambio* cuando se obtiene al intercambiar dos filas de I; y E es *de escala por r* si se obtiene al multiplicar una fila de I por un escalar r diferente de cero. Con esta terminología, el teorema 3 puede reformularse de la siguiente manera:

Si A es una matriz de $n \times n$ y E es una matriz elemental de $n \times n$, entonces

$$\det EA = (\det E)(\det A)$$

donde

$$\det E = \begin{cases} 1 & \textit{si } E \textit{ es un reemplazo de fila} \\ -1 & \textit{si } E \textit{ es de intercambio} \\ r & \textit{si } E \textit{ es de escala por } r \end{cases}$$

DEMOSTRACIÓN DEL TEOREMA 3 La demostración es por inducción sobre el tamaño de A. El caso de una matriz de 2×2 se verificó en los ejercicios 33 a 36 de la sección 3.1. Suponga que el teorema se ha verificado para determinantes de matrices de $k \times k$ con $k \geq 2$, sea $n = k + 1$, y sea A de $n \times n$.

En la acción de E sobre A intervienen ya sean dos filas o solamente una. Así que es posible desarrollar det EA a lo largo de una fila que no sea modificada por la acción de E, por ejemplo, la fila i. Sea A_{ij} (respectivamente, B_{ij}) la matriz obtenida al eliminar la fila i y la columna j de A (respectivamente, EA). Entonces las filas de B_{ij} se obtienen a partir de las filas de A_{ij} por medio de la misma operación elemental de fila que E realiza sobre A. Como estas submatrices son de solamente $k \times k$, la hipótesis de inducción implica que

$$\det B_{ij} = \alpha \cdot \det A_{ij}$$

donde $\alpha = 1, -1$ o r, dependiendo de la naturaleza de E. El desarrollo por cofactores a lo largo de la fila i es

$$\begin{aligned} \det EA &= a_{i1}(-1)^{i+1} \det B_{i1} + \cdots + a_{in}(-1)^{i+n} \det B_{in} \\ &= \alpha a_{i1}(-1)^{i+1} \det A_{i1} + \cdots + \alpha a_{in}(-1)^{i+n} \det A_{in} \\ &= \alpha \cdot \det A \end{aligned}$$

En particular, al considerar $A = I_n$, se observa que det $E = 1, -1$ o r, dependiendo de la naturaleza de E. Entonces el teorema es cierto para $n = 2$, y la validez del teorema para un valor de n implica que es cierto para el siguiente valor de n. Por el principio de inducción, el teorema debe ser cierto para $n \geq 2$. El teorema es trivialmente verdadero para $n = 1$. ■

DEMOSTRACIÓN DEL TEOREMA 6 Si A no es invertible, entonces AB tampoco lo es, por el ejercicio 27 de la sección 2.3. En este caso, det AB = (det A)(det B), porque ambos miembros son cero, de acuerdo con el teorema 4. Si A es invertible, entonces A y la matriz identidad I_n son equivalentes por filas, según el teorema de la matriz invertible. Así que existen matrices elementales $E_1, \ldots, E_p$ tales que

$$A = E_p E_{p-1} \cdots E_1 \cdot I_n = E_p E_{p-1} \cdots E_1$$

Por brevedad, se escribe $|A|$ en vez de det A. Entonces la aplicación repetida del teorema 3, tal como se replanteó anteriormente, muestra que

$$\begin{aligned} |AB| &= |E_p \cdots E_1 B| = |E_p||E_{p-1} \cdots E_1 B| = \cdots \\ &= |E_p| \cdots |E_1||B| = \cdots = |E_p \cdots E_1||B| \\ &= |A||B| \end{aligned}$$
■

PROBLEMAS DE PRÁCTICA

1. Calcule $\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix}$ en el menor número posible de pasos.

2. Use un determinante para decidir si $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ son linealmente independientes, cuando

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ -7 \\ 9 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 3 \\ -5 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -7 \\ 5 \end{bmatrix}$$

3.2 EJERCICIOS

Cada una de las ecuaciones que aparecen en los ejercicios 1 a 4 ilustra una propiedad de los determinantes. Enuncie la propiedad.

1. $\begin{vmatrix} 0 & 5 & -2 \\ 1 & -3 & 6 \\ 4 & -1 & 8 \end{vmatrix} = -\begin{vmatrix} 1 & -3 & 6 \\ 0 & 5 & -2 \\ 4 & -1 & 8 \end{vmatrix}$

2. $\begin{vmatrix} 2 & -6 & 4 \\ 3 & 5 & -2 \\ 1 & 6 & 3 \end{vmatrix} = 2\begin{vmatrix} 1 & -3 & 2 \\ 3 & 5 & -2 \\ 1 & 6 & 3 \end{vmatrix}$

3. $\begin{vmatrix} 1 & 3 & -4 \\ 2 & 0 & -3 \\ 5 & -4 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 3 & -4 \\ 0 & -6 & 5 \\ 5 & -4 & 7 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & -4 \\ 3 & 7 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & -4 \\ 0 & 1 & -5 \end{vmatrix}$

En los ejercicios 5 a 10, encuentre los determinantes mediante reducción por filas hasta la forma escalonada.

5. $\begin{vmatrix} 1 & 5 & -6 \\ -1 & -4 & 4 \\ -2 & -7 & 9 \end{vmatrix}$

6. $\begin{vmatrix} 1 & 5 & -3 \\ 3 & -3 & 3 \\ 2 & 13 & -7 \end{vmatrix}$

7. $\begin{vmatrix} 1 & 3 & 0 & 2 \\ -2 & -5 & 7 & 4 \\ 3 & 5 & 2 & 1 \\ 1 & -1 & 2 & -3 \end{vmatrix}$

8. $\begin{vmatrix} 1 & 3 & 3 & -4 \\ 0 & 1 & 2 & -5 \\ 2 & 5 & 4 & -3 \\ -3 & -7 & -5 & 2 \end{vmatrix}$

9. $\begin{vmatrix} 1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 2 & 8 & 5 \\ 3 & -1 & -2 & 3 \end{vmatrix}$

10. $\begin{vmatrix} 1 & 3 & -1 & 0 & -2 \\ 0 & 2 & -4 & -1 & -6 \\ -2 & -6 & 2 & 3 & 9 \\ 3 & 7 & -3 & 8 & -7 \\ 3 & 5 & 5 & 2 & 7 \end{vmatrix}$

Combine los métodos de reducción por filas y desarrollo por cofactores para calcular los determinantes en los ejercicios 11 a 14.

11. $\begin{vmatrix} 2 & 5 & -3 & -1 \\ 3 & 0 & 1 & -3 \\ -6 & 0 & -4 & 9 \\ 4 & 10 & -4 & -1 \end{vmatrix}$

12. $\begin{vmatrix} -1 & 2 & 3 & 0 \\ 3 & 4 & 3 & 0 \\ 5 & 4 & 6 & 6 \\ 4 & 2 & 4 & 3 \end{vmatrix}$

13. $\begin{vmatrix} 2 & 5 & 4 & 1 \\ 4 & 7 & 6 & 2 \\ 6 & -2 & -4 & 0 \\ -6 & 7 & 7 & 0 \end{vmatrix}$

14. $\begin{vmatrix} -3 & -2 & 1 & -4 \\ 1 & 3 & 0 & -3 \\ -3 & 4 & -2 & 8 \\ 3 & -4 & 0 & 4 \end{vmatrix}$

Encuentre los determinantes en los ejercicios 15 a 20, donde $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 7$.

15. $\begin{vmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{vmatrix}$

16. $\begin{vmatrix} a & b & c \\ 3d & 3e & 3f \\ g & h & i \end{vmatrix}$

17. $\begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix}$

18. $\begin{vmatrix} g & h & i \\ a & b & c \\ d & e & f \end{vmatrix}$

19. $\begin{vmatrix} a & b & c \\ 2d+a & 2e+b & 2f+c \\ g & h & i \end{vmatrix}$

20. $\begin{vmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{vmatrix}$

En los ejercicios 21 a 23, utilice determinantes para averiguar si la matriz es invertible.

21. $\begin{bmatrix} 2 & 3 & 0 \\ 1 & 3 & 4 \\ 1 & 2 & 1 \end{bmatrix}$

22. $\begin{bmatrix} 5 & 0 & -1 \\ 1 & -3 & -2 \\ 0 & 5 & 3 \end{bmatrix}$

23. $\begin{bmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{bmatrix}$

En los ejercicios 24 a 26, utilice determinantes para averiguar si el conjunto de vectores es linealmente independiente.

24. $\begin{bmatrix} 4 \\ 6 \\ -7 \end{bmatrix}, \begin{bmatrix} -7 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ -5 \\ 6 \end{bmatrix}$

25. $\begin{bmatrix} 7 \\ -4 \\ -6 \end{bmatrix}, \begin{bmatrix} -8 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ -5 \end{bmatrix}$

26. $\begin{bmatrix} 3 \\ 5 \\ -6 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -6 \\ 0 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -3 \end{bmatrix}$

En los ejercicios 27 y 28, A y B son matrices de $n \times n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

27. a. Una operación de reemplazo de filas no afecta el determinante de una matriz.

b. El determinante de A es el producto de los pivotes presentes en cualquier forma escalonada U de A, multiplicado por $(-1)^r$, donde r es el número de intercambios de fila realizados durante la reducción por filas de A a U.

c. Si las columnas de A son linealmente dependientes, entonces $\det A = 0$.

d. $\det(A + B) = \det A + \det B$.

28. a. Si se realizan dos intercambios sucesivos de fila, entonces el nuevo determinante es igual al determinante antiguo.

b. El determinante de A es el producto de las entradas diagonales de A.

c. Si $\det A$ es cero, entonces dos filas o dos columnas son iguales, o una fila o una columna es cero.

d. $\det A^T = (-1) \det A$.

29. Calcule $\det B^5$, donde $B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$.

30. Use el teorema 3 (pero no el teorema 4) para demostrar que si dos filas de una matriz cuadrada A son iguales, entonces $\det A = 0$. Esto se cumple también para dos columnas. ¿Por qué?

En los ejercicios 31 a 36, mencione en la explicación un teorema apropiado.

31. Muestre que si A es invertible, entonces $\det A^{-1} = \dfrac{1}{\det A}$.

32. Encuentre una fórmula para $\det(rA)$ cuando A es una matriz de $n \times n$.

33. Sean A y B matrices cuadradas. Muestre que aunque AB y BA no sean iguales, siempre es cierto que $\det AB = \det BA$.

34. Sean A y P matrices cuadradas, con P invertible. Muestre que $\det(PAP^{-1}) = \det A$.

35. Sea U una matriz cuadrada tal que $U^TU = I$. Muestre que $\det U = \pm 1$.

36. Suponga que A es una matriz cuadrada tal que $\det A^4 = 0$. Explique por qué A no puede ser invertible.

Verifique que $\det AB = (\det A)(\det B)$ para las matrices de los ejercicios 37 y 38. (No utilice el teorema 6.)

37. $A = \begin{bmatrix} 3 & 0 \\ 6 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 5 & 4 \end{bmatrix}$

38. $A = \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 2 \\ -1 & -1 \end{bmatrix}$

39. Sean A y B matrices de 3×3, con $\det A = 4$ y $\det B = -3$. Utilice las propiedades de los determinantes (dadas en el texto y en los ejercicios anteriores) para calcular:

a. $\det AB$ b. $\det 5A$ c. $\det B^T$

d. $\det A^{-1}$ e. $\det A^3$

40. Sean A y B matrices de 4×4, con $\det A = -1$ y $\det B = 2$. Calcule:

a. $\det AB$ b. $\det B^5$ c. $\det 2A$

d. $\det A^TA$ e. $\det B^{-1}AB$

41. Verifique que $A = \det B + \det C$, donde

$$A = \begin{bmatrix} a+e & b+f \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, C = \begin{bmatrix} e & f \\ c & d \end{bmatrix}$$

42. Sean $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Muestre que $\det(A + B) = \det A + \det B$ si, y sólo si, $a + d = 0$.

43. Muestre que $\det A = \det B + \det C$, donde

$$A = \begin{bmatrix} a_{11} & a_{12} & u_1 + v_1 \\ a_{21} & a_{22} & u_2 + v_2 \\ a_{31} & a_{32} & u_3 + v_3 \end{bmatrix},$$

$$B = \begin{bmatrix} a_{11} & a_{12} & u_1 \\ a_{21} & a_{22} & u_2 \\ a_{31} & a_{32} & u_3 \end{bmatrix}, \quad C = \begin{bmatrix} a_{11} & a_{12} & v_1 \\ a_{21} & a_{22} & v_2 \\ a_{31} & a_{32} & v_3 \end{bmatrix}$$

Sin embargo, observe que A *no* es lo mismo que $B + C$.

44. La multiplicación derecha por una matriz elemental E afecta las *columnas* de A en la misma forma que la multiplicación izquierda afecta las *filas*. Utilice los teoremas 5 y 3, y el hecho evidente de que E^T es otra matriz elemental, para mostrar que

$\det AE = (\det E)(\det A)$

No use el teorema 6.

45. **[M]** Calcule $\det A^TA$ y $\det AA^T$ para varias matrices aleatorias de 4×5 y varias matrices aleatorias de 5×6. ¿Qué puede decirse acerca de A^TA y de AA^T cuando A tiene más columnas que filas?

46. **[M]** Si $\det A$ es cercano a cero, ¿la matriz A es casi singular? Experimente con la matriz casi triangular A de 4×4 del ejercicio 9 presentado en la sección 2.3. Encuentre los determinantes de A, $10A$ y $0.1A$. En contraste, determine los números de condición de estas matrices. Repita estos cálculos cuando A es la matriz identidad de 4×4. Analice sus resultados.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Efectúe reemplazos de filas para crear ceros en la primera columna, y luego genere una fila de ceros.

$$\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 1 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} = 0$$

2. $$\det[\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] = \begin{vmatrix} 5 & -3 & 2 \\ -7 & 3 & -7 \\ 9 & -5 & 5 \end{vmatrix} = \begin{vmatrix} 5 & -3 & 2 \\ -2 & 0 & -5 \\ 9 & -5 & 5 \end{vmatrix}$$ Fila 1 sumada a la fila 2

$$= -(-3)\begin{vmatrix} -2 & -5 \\ 9 & 5 \end{vmatrix} - (-5)\begin{vmatrix} 5 & 2 \\ -2 & -5 \end{vmatrix}$$ Cofactores de la columna 2

$$= 3 \cdot (35) + 5 \cdot (-21) = 0$$

Por el teorema 4, la matriz $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$ no es invertible. Las columnas son linealmente dependientes, de acuerdo con el teorema de la matriz invertible.

3.3 REGLA DE CRAMER, VOLUMEN Y TRANSFORMACIONES LINEALES

Esta sección aplica la teoría de las secciones anteriores para obtener fórmulas teóricas de gran importancia y una interpretación geométrica de los determinantes.

Regla de Cramer

La regla de Cramer se necesita en diversos cálculos teóricos. Por ejemplo, puede utilizarse para estudiar cómo se modifica la solución de $A\mathbf{x} = \mathbf{b}$ cuando cambian las entradas de $\mathbf{b}$. Sin embargo, la fórmula es ineficiente para cálculos a mano, excepto para matrices de 2×2 o, quizá, de 3×3.

Para toda matriz A de $n \times n$ y cualquier $\mathbf{b}$ en $\mathbb{R}^n$, sea $A_i(\mathbf{b})$ la matriz obtenida a partir de A mediante el reemplazo de la columna i por el vector $\mathbf{b}$.

$$A_i(\mathbf{b}) = [\mathbf{a}_1 \quad \cdots \quad \underset{\substack{\uparrow \\ \text{col } i}}{\mathbf{b}} \quad \cdots \quad \mathbf{a}_n]$$

TEOREMA 7 Regla de Cramer

Sea A una matriz invertible $n \times n$. Para cualquier $\mathbf{b}$ en $\mathbb{R}^n$, la solución única $\mathbf{x}$ de $A\mathbf{x} = \mathbf{b}$ tiene entradas dadas por

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \qquad i = 1, 2, \ldots, n \tag{1}$$

DEMOSTRACIÓN Denote las columnas de A mediante $\mathbf{a}_1, \dots, \mathbf{a}_n$ y las columnas de la matriz identidad I de $n \times n$ por medio de $\mathbf{e}_1, \dots, \mathbf{e}_n$. Si $A\mathbf{x} = \mathbf{b}$, la definición de multiplicación de matrices muestra que

$$\begin{aligned} A \cdot I_i(\mathbf{x}) &= A[\,\mathbf{e}_1 \quad \cdots \quad \mathbf{x} \quad \cdots \quad \mathbf{e}_n\,] = [\,A\mathbf{e}_1 \quad \cdots \quad A\mathbf{x} \quad \cdots \quad A\mathbf{e}_n\,] \\ &= [\,\mathbf{a}_1 \quad \cdots \quad \mathbf{b} \quad \cdots \quad \mathbf{a}_n\,] = A_i(\mathbf{b}) \end{aligned}$$

Por la propiedad multiplicativa de los determinantes,

$$(\det A)(\det I_i(\mathbf{x})) = \det A_i(\mathbf{b})$$

El segundo determinante de la izquierda es simplemente x_i. (Efectúe un desarrollo por cofactores a lo largo de la i-ésima fila.) Por lo tanto $(\det A) \cdot x_i = \det A_i(\mathbf{b})$. Esto demuestra (1) dado que A es invertible y $\det A \neq 0$. ■

EJEMPLO 1 Use la regla de Cramer para resolver el sistema

$$\begin{aligned} 3x_1 - 2x_2 &= 6 \\ -5x_1 + 4x_2 &= 8 \end{aligned}$$

Solución Vea el sistema como $A\mathbf{x} = \mathbf{b}$. Usando la notación que se introdujo anteriormente,

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, \qquad A_1(\mathbf{b}) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, \qquad A_2(\mathbf{b}) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

Como $\det A = 2$, el sistema tiene una solución única. Por la regla de Cramer,

$$\begin{aligned} x_1 &= \frac{\det A_1(\mathbf{b})}{\det A} = \frac{24 + 16}{2} = 20 \\ x_2 &= \frac{\det A_2(\mathbf{b})}{\det A} = \frac{24 + 30}{2} = 27 \end{aligned}$$

■

Aplicación a la ingeniería

Muchos problemas importantes en ingeniería, particularmente en ingeniería eléctrica y en teoría de control, se pueden analizar por medio de *transformaciones de Laplace.* Este enfoque convierte un sistema adecuado de ecuaciones diferenciales lineales en un sistema de ecuaciones algebraicas lineales cuyos coeficientes incluyen un parámetro. El siguiente ejemplo ilustra el tipo de sistema algebraico que puede surgir.

EJEMPLO 2 Considere el siguiente sistema, en el cual s es un parámetro no especificado. Determine los valores de s para los cuales el sistema tiene una solución única, y use la regla de Cramer para describir la solución.

$$\begin{aligned} 3sx_1 - 2x_2 &= 4 \\ -6x_1 + sx_2 &= 1 \end{aligned}$$

Solución Vea el sistema como $A\mathbf{x} = \mathbf{b}$. Entonces

$$A = \begin{bmatrix} 3s & -2 \\ -6 & s \end{bmatrix}, \qquad A_1(\mathbf{b}) = \begin{bmatrix} 4 & -2 \\ 1 & s \end{bmatrix}, \qquad A_2(\mathbf{b}) = \begin{bmatrix} 3s & 4 \\ -6 & 1 \end{bmatrix}$$

Dado que

$$\det A = 3s^2 - 12 = 3(s+2)(s-2)$$

el sistema tiene una solución única precisamente cuando $s \neq \pm 2$. Para una s como ésta, la solución es (x_1, x_2), donde

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{4s+2}{3(s+2)(s-2)}$$

$$x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{3s+24}{3(s+2)(s-2)} = \frac{s+8}{(s+2)(s-2)}$$

Una fórmula para A^{-1}

La regla de Cramer conduce fácilmente a una fórmula general para el inverso de una matriz A $n \times n$. La j-ésima columna de A^{-1} es un vector $\mathbf{x}$ que satisface

$$A\mathbf{x} = \mathbf{e}_j$$

donde $\mathbf{e}_j$ es la j-ésima columna de la matriz identidad, y la i-ésima entrada de $\mathbf{x}$ es la entrada (i, j) de A^{-1}. Por la regla de Cramer,

$$\left\{\text{entrada } (i,\ j) \text{ de } A^{-1}\right\} = x_i = \frac{\det A_i(\mathbf{e}_j)}{\det A} \tag{2}$$

Recuerde que A_{ji} denota a la submatriz de A que se forma al eliminar la fila j y la columna i. Un desarrollo por cofactores descendiendo por la columna i de $A_i(\mathbf{e}_j)$ muestra que

$$\det A_i(\mathbf{e}_j) = (-1)^{i+j} \det A_{ji} = C_{ji} \tag{3}$$

donde C_{ji} es un cofactor de A. Mediante (2), la entrada (i, j) de A^{-1} es el cofactor C_{ji} dividido entre $\det A$. [Observe que los subíndices de C_{ji} son el inverso de (i, j).] Entonces

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \tag{4}$$

La matriz de cofactores del miembro derecho de (4) es la **adjunta** (o **adjunta clásica**) de A, denotada mediante adj A. (El término *adjunta* tiene también otro significado en los textos sobre transformaciones lineales.) El teorema presentado enseguida simplemente replantea (4).

TEOREMA 8 Una fórmula para el inverso

Sea A una matriz invertible de $n \times n$. Entonces

$$A^{-1} = \frac{1}{\det A}\,\text{adj}\,A$$

EJEMPLO 3 Encuentre el inverso de la matriz $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$.

Solución Los nueve cofactores son

$$C_{11} = +\begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2, \quad C_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3, \quad C_{13} = +\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14, \quad C_{22} = +\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7, \quad C_{23} = -\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = +\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4, \quad C_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1, \quad C_{33} = +\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

La matriz adjunta es la *transpuesta* de la matriz de cofactores. [Por ejemplo, C_{12} va en la posición (2, 1).] Entonces

$$\text{adj}\, A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

Se podría encontrar det A en forma directa, pero el siguiente cálculo verifica los cálculos anteriores *y* además produce det A:

$$(\text{adj}\, A)\cdot A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

Como (adj A) $A = 14I$, el teorema 8 muestra que det $A = 14$, y

$$A^{-1} = \frac{1}{14}\begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} = \begin{bmatrix} -1/7 & 1 & 2/7 \\ 3/14 & -1/2 & 1/14 \\ 5/14 & -1/2 & -3/14 \end{bmatrix}$$

NOTAS NUMÉRICAS

El teorema 8 resulta útil, principalmente, para efectuar los cálculos teóricos. La fórmula de A^{-1} permite deducir propiedades del inverso sin tener que calcularlo. Excepto en casos especiales, el algoritmo de la sección 2.2 proporciona una mucho mejor manera de calcular A^{-1}, si el inverso es realmente necesario.

La regla de Cramer también es una herramienta teórica. Puede servir para estudiar qué tan sensibles son las soluciones de $A\mathbf{x} = \mathbf{b}$ a los cambios efectuados en una entrada de $\mathbf{b}$ o de A (cambios debidos quizá al error experimental cuando se obtienen las entradas para $\mathbf{b}$ o A). Cuando A es una matriz de 3×3 con entradas *complejas*, ocasionalmente se elige la regla de Cramer para realizar cálculos a mano porque la reducción por filas de $[A \;\; \mathbf{b}]$ mediante aritmética compleja puede resultar complicada, y los determinantes son relativamente fáciles de calcular. Para una matriz de $n \times n$ más grande (real o compleja), la regla de Cramer resulta, sin duda, ineficiente. Calcular tan sólo *un* determinante requiere tanto trabajo como resolver mediante reducción por filas $A\mathbf{x} = \mathbf{b}$.

Determinantes como área o volumen

En la siguiente aplicación se verificará la interpretación geométrica de los determinantes descrita en la introducción del capítulo. Aunque no se efectuará un análisis general de

longitud y distancia en $\mathbb{R}^n$ sino hasta el capítulo 6, aquí se supone que los conceptos euclidianos usuales de longitud, área y volumen para $\mathbb{R}^2$ y $\mathbb{R}^3$ ya son comprendidos.

TEOREMA 9

Si A es una matriz de 2×2, el área del paralelogramo determinado por las columnas de A es $|\det A|$. Si A es una matriz de 3×3, el volumen del paralelepípedo determinado mediante las columnas de A es $|\det A|$.

SG Una demostración geométrica 3 a 12 (A Geometric Proof 3-12)

DEMOSTRACIÓN Es evidente que el teorema resulta cierto para cualquier matriz diagonal:

$$\left|\det \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}\right| = |ad| = \left\{\begin{matrix}\text{área del} \\ \text{rectángulo}\end{matrix}\right\}$$

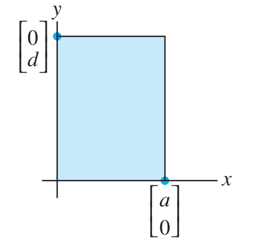


FIGURA 1
Área $= |ad|$.

Vea la figura 1. Es suficiente con demostrar que cualquier matriz de 2×2 $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$ se puede transformar en una matriz diagonal de manera que no cambie ni el área del paralelogramo asociado ni $|\det A|$. De la sección 3.2, se sabe que el valor absoluto del determinante no cambia cuando se intercambian dos columnas o se suma una fila con el múltiplo de otra. Y es fácil advertir que dichas operaciones bastan para transformar A en una matriz diagonal. Los intercambios de columna no modifican en modo alguno al paralelogramo. Así que es suficiente con demostrar la sencilla observación geométrica siguiente que se aplica a los vectores en $\mathbb{R}^2$ o $\mathbb{R}^3$:

Sean $\mathbf{a}_1$ y $\mathbf{a}_2$ vectores diferentes de cero. Entonces, para cualquier escalar c, el área del paralelogramo determinado mediante $\mathbf{a}_1$ y $\mathbf{a}_2$ es igual al área del paralelogramo determinado por $\mathbf{a}_1$ y $\mathbf{a}_2 + c\mathbf{a}_1$.

Para demostrar este enunciado, puede suponerse que $\mathbf{a}_2$ no es un múltiplo de $\mathbf{a}_1$, porque de serlo los dos paralelogramos serían degenerados y tendrían área cero. Si L es la línea que pasa por $\mathbf{0}$ y $\mathbf{a}_1$, entonces $\mathbf{a}_2 + L$ es la línea paralela a L que pasa por $\mathbf{a}_2$, y $\mathbf{a}_2 + c\mathbf{a}_1$ está sobre esta línea. Vea la figura 2. Los puntos $\mathbf{a}_2$ y $\mathbf{a}_2 + c\mathbf{a}_1$ tienen la misma distancia perpendicular a L. Por lo tanto, los dos paralelogramos de la figura 2 tienen la misma área, ya que comparten la base de $\mathbf{0}$ a $\mathbf{a}_1$. Esto completa la comprobación para $\mathbb{R}^2$.

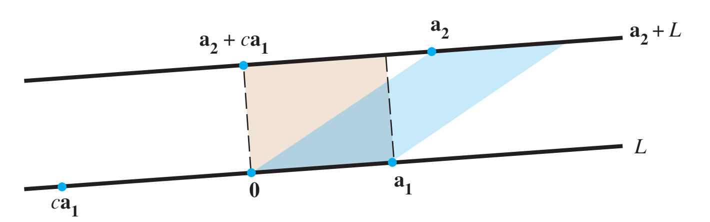


FIGURA 2 Dos paralelogramos de igual área.

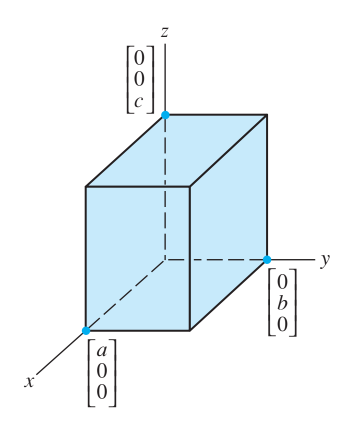


FIGURA 3
Volumen $= |abc|$.

La demostración para $\mathbb{R}^3$ es similar. Resulta evidente que el teorema es válido para una matriz diagonal de 3×3. Vea la figura 3. Y cualquier matriz A de 3×3 se puede transformar en una matriz diagonal con operaciones por columna que no modifican $|\det A|$. (Piense en hacer operaciones por fila con A^T.) Por lo tanto, es suficiente con demostrar que estas operaciones no afectan el volumen del paralelepípedo determinado por las columnas de A.

En la figura 4 se muestra un paralelepípedo en forma de una caja sombreada con dos lados inclinados. Su volumen es el área de la base en el plano Gen$\{\mathbf{a}_1, \mathbf{a}_3)$ por la altura de $\mathbf{a}_2$ sobre Gen$\{\mathbf{a}_1, \mathbf{a}_3\}$. Cualquier vector $\mathbf{a}_2 + c\mathbf{a}_1$ tiene la misma altura porque $\mathbf{a}_2 + c\mathbf{a}_1$ está en el plano $\mathbf{a}_2 +$ Gen$\{\mathbf{a}_1, \mathbf{a}_3\}$, el cual es paralelo a Gen$\{\mathbf{a}_1, \mathbf{a}_3\}$. Por lo tanto, no cambia el volumen del paralelepípedo cuando $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$ cambia a $[\mathbf{a}_1 \quad \mathbf{a}_2 + c\mathbf{a}_1 \quad \mathbf{a}_3]$. Entonces, una operación de reemplazo de columna no afecta el volumen del paralelepípedo. Puesto que los intercambios de columna no tienen efecto sobre el volumen, la demostración está completa. ■

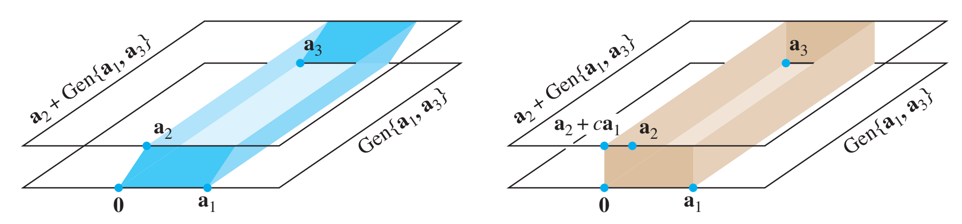


FIGURA 4 Dos paralelepípedos de igual volumen.

EJEMPLO 4 Calcule el área del paralelogramo determinado por los puntos $(-2, -2)$, $(0, 3)$, $(4, -1)$ y $(6, 4)$. Vea la figura 5(a).

Solución Primero traslade el paralelogramo a uno que tenga el origen como vértice. Por ejemplo, reste el vértice $(-2, -2)$ de cada uno de los cuatro vértices. El nuevo paralelogramo tiene la misma área, y sus vértices son $(0, 0)$, $(2, 5)$, $(6, 1)$ y $(8, 6)$. Vea la figura 5(b). Este paralelogramo está determinado por las columnas de

$$A = \begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}$$

Como $|\det A| = |-28|$, el área del paralelogramo es 28. ■

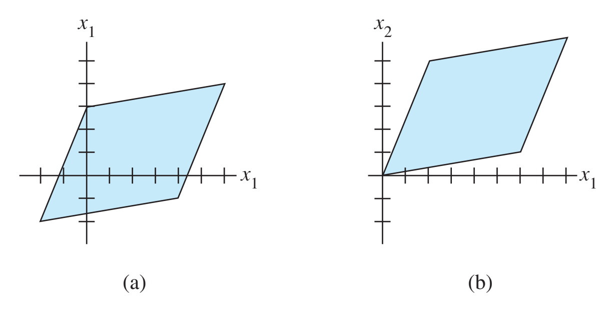


FIGURA 5 La traslación de un paralelogramo no altera su área.

Transformaciones lineales

Los determinantes pueden usarse para describir una importante propiedad geométrica de las transformaciones lineales en el plano y en $\mathbb{R}^3$. Si T es una transformación lineal y S un conjunto en el dominio de T, denote con $T(S)$ al conjunto de imágenes de puntos localizados en S. Se desea comparar el área (o volumen) de $T(S)$ con el área (o volumen) del conjunto original S. Por convención, cuando S es una región delimitada por un paralelogramo, se hace referencia a S como un paralelogramo.

TEOREMA 10

Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal determinada mediante una matriz A. Si S es un paralelogramo en $\mathbb{R}^2$, entonces

$$\{\text{área de } T(S)\} = |\det A| \cdot \{\text{área de } S\} \tag{5}$$

Si T está determinada por una matriz A de 3×3, y si S es un paralelepípedo en $\mathbb{R}^3$, entonces

$$\{\text{volumen de } T(S)\} = |\det A| \cdot \{\text{volumen de } S\} \tag{6}$$

DEMOSTRACIÓN Considere el caso de 2×2, con $A = \{\mathbf{a}_1 \quad \mathbf{a}_2\}$. Un paralelogramo en el origen en $\mathbb{R}^2$ determinado por vectores $\mathbf{b}_1$ y $\mathbf{b}_2$ tiene la forma

$$S = \{s_1\mathbf{b}_1 + s_2\mathbf{b}_2 : 0 \leq s_1 \leq 1,\ 0 \leq s_2 \leq 1\}$$

La imagen de S bajo T consiste en puntos de la forma

$$\begin{aligned} T(s_1\mathbf{b}_1 + s_2\mathbf{b}_2) &= s_1T(\mathbf{b}_1) + s_2T(\mathbf{b}_2) \\ &= s_1A\mathbf{b}_1 + s_2A\mathbf{b}_2 \end{aligned}$$

donde $0 \leq s_1 \leq 1$, $0 \leq s_2 \leq 1$. Se sigue que $T(S)$ es el paralelogramo determinado por las columnas de la matriz $[A\mathbf{b}_1 \quad A\mathbf{b}_2]$. Esta matriz puede escribirse como AB, donde $B = [\mathbf{b}_1 \quad \mathbf{b}_2]$. De acuerdo con el teorema 9 y el teorema del producto para determinantes,

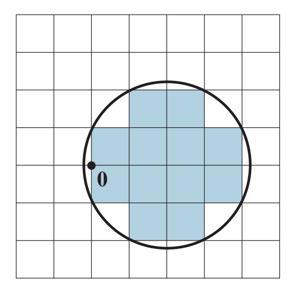
$$\begin{aligned} \{\text{área de } T(S)\} &= |\det AB| = |\det A| \cdot |\det B| \\ &= |\det A| \cdot \{\text{área de } S\} \end{aligned} \tag{7}$$

Un paralelogramo arbitrario tiene la forma $\mathbf{p} + S$, donde $\mathbf{p}$ es un vector y S un paralelogramo en el origen, como el anterior. Resulta fácil advertir que T transforma a $\mathbf{p} + S$ en $T(\mathbf{p}) + T(S)$. (Vea el ejercicio 26.) Como la traslación no afecta el área de un conjunto,

$$\begin{aligned} \{\text{área de } T(\mathbf{p} + S)\} &= \{\text{área de } T(\mathbf{p}) + T(S)\} \\ &= \{\text{área de } T(S)\} && \text{Traslación} \\ &= |\det A| \cdot \{\text{área de } S\} && \text{Por (7)} \\ &= |\det A| \cdot \{\text{área de } \mathbf{p} + S\} && \text{Traslación} \end{aligned}$$

Esto demuestra que (5) es válida para todos los paralelogramos en $\mathbb{R}^2$. La demostración de (6) para el caso de 3×3 es análoga. ■

Cuando se intenta generalizar el teorema 10 a una región en $\mathbb{R}^2$ o en $\mathbb{R}^3$ que no está delimitada por líneas rectas o planos, debe afrontarse el problema de cómo definir y calcular su área o volumen. Éste es un tema que se estudia en cálculo, ahora simplemente se delineará la idea básica para $\mathbb{R}^2$. Si R es una región plana que tiene área finita, entonces puede aproximarse mediante una malla de pequeños cuadrados que estén dentro de R. Si los cuadrados se hacen lo suficientemente pequeños, el área de R puede aproximarse tanto como se desee al sumar las áreas de esos pequeños cuadrados. Vea la figura 6.



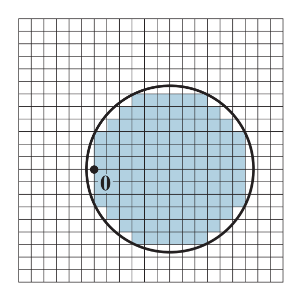


FIGURA 6 Aproximación de una región plana mediante cuadrados. La aproximación mejora al hacerse más fina la cuadrícula.

Si T es una transformación lineal asociada con una matriz A de 2×2, entonces la imagen de una región plana R bajo T se aproxima mediante las imágenes de los pequeños cuadrados trazados dentro de R. La prueba del teorema 10 muestra que toda imagen de este tipo es un paralelogramo cuya área es $|\det A|$ multiplicado por el área del cuadrado. Si R' es la unión de los cuadrados trazados dentro de R, entonces el área de $T(R')$ es $|\det A|$ multiplicado por el área de R'. Vea la figura 7. También, el área de $T(R')$ es cercana al área de $T(R)$. Puede darse un argumento que use un proceso restrictivo para justificar la siguiente generalización del teorema 10.

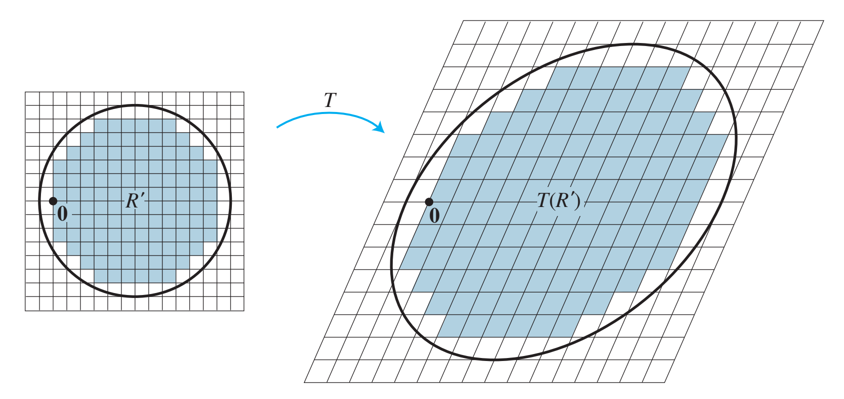
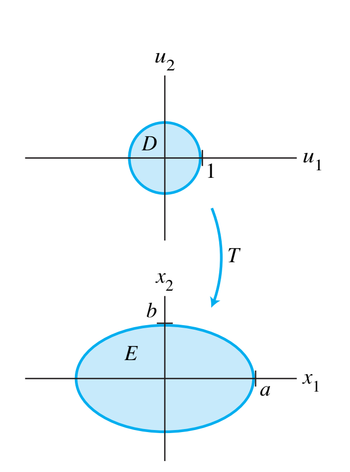


FIGURA 7 Aproximación de $T(R)$ mediante una unión de paralelogramos.

Las conclusiones del teorema 10 son válidas siempre que S sea una región de $\mathbb{R}^2$ con área finita, o en una región de $\mathbb{R}^3$ con volumen finito.

EJEMPLO 5 Sean a y b números positivos. Encuentre el área de la región E delimitada por la elipse cuya ecuación es

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$



Solución Se afirma que E es la imagen del disco unitario D bajo la transformación lineal T determinada mediante la matriz $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, porque si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ y $\mathbf{x} = A\mathbf{u}$, entonces

$$u_1 = \frac{x_1}{a} \qquad \text{y} \qquad u_2 = \frac{x_2}{b}$$

Se sigue que $\mathbf{u}$ está en el disco unitario, con $u_1^2 + u_2^2 \leq 1$, si, y sólo si, $\mathbf{x}$ está en E, con $(x_1/a)^2 + (x_2/b)^2 \leq 1$. Por la generalización del teorema 10,

$$\begin{aligned} \{\text{área del elipse}\} &= \{\text{área de } T(D)\} \\ &= |\det A| \cdot \{\text{área de } D\} \\ &= ab \cdot \pi(1)^2 = \pi ab \end{aligned}$$

PROBLEMA DE PRÁCTICA

Sea S el paralelogramo determinado mediante los vectores $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, y $A = \begin{bmatrix} 1 & -.1 \\ 0 & 2 \end{bmatrix}$. Calcule el área de la imagen de S bajo la función $\mathbf{x} \mapsto A\mathbf{x}$.

3.3 EJERCICIOS

Use la regla de Cramer para calcular las soluciones a los sistemas de los ejercicios 1 a 6.

1. $\begin{aligned} 5x_1 + 7x_2 &= 3 \\ 2x_1 + 4x_2 &= 1 \end{aligned}$

2. $\begin{aligned} 4x_1 + x_2 &= 6 \\ 5x_1 + 2x_2 &= 7 \end{aligned}$

3. $\begin{aligned} 3x_1 - 2x_2 &= 7 \\ -5x_1 + 6x_2 &= -5 \end{aligned}$

4. $\begin{aligned} -5x_1 + 3x_2 &= 9 \\ 3x_1 - x_2 &= -5 \end{aligned}$

5. $\begin{aligned} 2x_1 + x_2 &= 7 \\ -3x_1 + x_3 &= -8 \\ x_2 + 2x_3 &= -3 \end{aligned}$

6. $\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ -x_1 + 2x_3 &= 2 \\ 3x_1 + x_2 + 3x_3 &= -2 \end{aligned}$

En los ejercicios 7 a 10, determine los valores del parámetro S para los cuales el sistema tiene una solución única y describa dicha solución.

7. $\begin{aligned} 6sx_1 + 4x_2 &= 5 \\ 9x_1 + 2sx_2 &= -2 \end{aligned}$

8. $\begin{aligned} 3sx_1 - 5x_2 &= 3 \\ 9x_1 + 5sx_2 &= 2 \end{aligned}$

9. $\begin{aligned} sx_1 - 2sx_2 &= -1 \\ 3x_1 + 6sx_2 &= 4 \end{aligned}$

10. $\begin{aligned} 2sx_1 + x_2 &= 1 \\ 3sx_1 + 6sx_2 &= 2 \end{aligned}$

En los ejercicios 11 a 16, calcule la adjunta de la matriz dada, y luego utilice el teorema 8 para encontrar el inverso de la matriz.

11. $\begin{bmatrix} 0 & -2 & -1 \\ 3 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 1 & 3 \\ 2 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

13. $\begin{bmatrix} 3 & 5 & 4 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 3 & 6 & 7 \\ 0 & 2 & 1 \\ 2 & 3 & 4 \end{bmatrix}$

15. $\begin{bmatrix} 3 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 3 & 2 \end{bmatrix}$

16. $\begin{bmatrix} 1 & 2 & 4 \\ 0 & -3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$

17. Muestre que si A es de 2×2, entonces el teorema 8 proporciona la misma fórmula para A^{-1} que la dada por el teorema 4 en la sección 2.2.

18. Suponga que todas las entradas en A son enteros y que $\det A = 1$. Explique por qué todas las entradas de A^{-1} son enteros.

En los ejercicios 19 a 22, encuentre el área del paralelogramo cuyos vértices son los que se enlistan.

19. (0, 0), (5, 2), (6, 4), (11, 6)

20. (0, 0), (−1, 3), (4, −5), (3, −2)

21. (−1, 0), (0, 5), (1, −4), (2, 1)

22. (0, −2), (6, −1), (−3, 1), (3, 2)

23. Encuentre el volumen del paralelepípedo que tiene un vértice en el origen y vértices adyacentes en (1, 0, −2), (1, 2, 4), (7, 1, 0).

24. Encuentre el volumen del paralelepípedo que tiene un vértice en el origen y vértices adyacentes en (1, 4, 0), (−2, −5, 2), (−1, 2, −1).

25. Use el concepto de volumen para explicar por qué el determinante de una matriz A de 3×3 es cero si, y sólo si, A no es invertible. No recurra al teorema 4 de la sección 3.2. [*Sugerencia:* Piense en las columnas de A.]

26. Sea $T : \mathbb{R}^m \to \mathbb{R}^n$ una transformación lineal, y sean $\mathbf{p}$ un vector y S un conjunto en $\mathbb{R}^m$. Muestre que la imagen de $\mathbf{p} + S$ bajo T es el conjunto trasladado $T(\mathbf{p}) + T(S)$ en $\mathbb{R}^n$.

27. Sea S el paralelogramo determinado por los vectores $\mathbf{b}_1 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$, y sea $A = \begin{bmatrix} 6 & -2 \\ -3 & 2 \end{bmatrix}$. Calcule el área de la imagen de S bajo la función $\mathbf{x} \mapsto A\mathbf{x}$.

28. Repita el ejercicio 27 con $\mathbf{b}_1 = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, y $A = \begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix}$.

29. Encuentre una fórmula para el área del triángulo cuyos vértices son $\mathbf{0}$, $\mathbf{v}_1$ y $\mathbf{v}_2$ en $\mathbb{R}^2$.

30. Sea R el triángulo con vértices en (x_1, y_1), (x_2, y_2), y (x_3, y_3). Muestre que

$$\{\text{área del triángulo}\} = \frac{1}{2}\left|\det \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}\right|$$

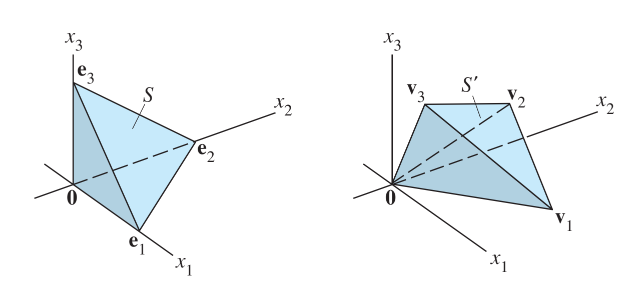
[*Sugerencia:* Traslade R al origen restando uno de los vértices, y use el ejercicio 29.]

31. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal determinada mediante la matriz $A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, donde a, b y c son números positivos. Sea S la esfera unitaria, cuya superficie limitante tiene la ecuación $x_1^2 + x_2^2 + x_3^2 = 1$.

a. Muestre que $T(S)$ está delimitada por el elipsoide que tiene la ecuación $\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} + \dfrac{x_3^2}{c^2} = 1$.

b. Utilice el hecho de que el volumen de la esfera unitaria es $4\pi/3$ para determinar el volumen de la región acotada por el elipsoide del inciso (a).

32. Sea S el tetraedro en $\mathbb{R}^3$ con vértices en los vectores $\mathbf{0}$, $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$, y sea S' el tetraedro con vértices en los vectores $\mathbf{0}$, $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Vea la siguiente figura.



a. Describa una transformación lineal que mapee S sobre S'.

b. Encuentre una fórmula para el volumen del tetraedro S' utilizando el hecho de que

$\{\text{volumen de } S\} = (1/3)\{\text{área de la base}\}\cdot\{\text{altura}\}$.

33. **[M]** Pruebe la fórmula del inverso del teorema 8 para una matriz arbitraria A de 4×4. Use un programa de matrices para calcular los cofactores de las submatrices 3×3, construya la adjunta, y establezca $B = (\text{adj } A)/(\det A)$. Luego calcule $B - \text{inv}(A)$, donde $\text{inv}(A)$ es el inverso de A calculado mediante el programa de matrices. Utilice aritmética de punto flotante con el máximo posible de posiciones decimales. Informe acerca de sus resultados.

34. **[M]** Pruebe la regla de Cramer para una matriz aleatoria A de 4×4 y un vector $\mathbf{b}$ aleatorio de 4×1. En la solución de $A\mathbf{x} = \mathbf{b}$, calcule cada entrada y compare esas entradas con las entradas de $A^{-1}\mathbf{b}$. Escriba el comando (o golpes de tecla) para el programa de computadora que usa la regla de Cramer para producir la segunda entrada de $\mathbf{x}$.

35. **[M]** Si su versión de MATLAB tiene el comando `flops`, úselo para contar el número de operaciones de punto flotante requeridas para calcular el inverso de una matriz aleatoria de 30×30. Compare este número con los flops necesarios para formar $(\text{adj } A)/(\det A)$.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

El área de S es $\left|\det \begin{bmatrix} 1 & 5 \\ 3 & 1 \end{bmatrix}\right| = 14$, y $\det A = 2$. Por el teorema 10, el área de la imagen de S bajo la función $\mathbf{x} \mapsto A\mathbf{x}$ es

$$|\mathbf{det}\, A| \cdot \{\text{área de } S\} = 2 \cdot 14 = 28$$

CAPÍTULO 3 EJERCICIOS SUPLEMENTARIOS

1. Señale cada enunciado como verdadero o falso. Justifique sus respuestas. Suponga que todas las matrices son cuadradas.

a. Si A es una matriz de 2×2 con determinante cero, entonces una columna de A es múltiplo de la otra columna.

b. Si dos filas de una matriz A de 3×3 son iguales, entonces $\det A = 0$.

c. Si A es una matriz de 3×3, entonces $\det 5A = 5 \det A$.

d. Si A y B son matrices de $n \times n$, con $\det A = 2$ y $\det B = 3$, entonces $\det(A + B) = 5$.

e. Si A es de $n \times n$ y $\det A = 2$, entonces $\det A^3 = 6$.

f. Si B se produce al intercambiar dos filas de A, entonces $\det B = \det A$.

g. Si B se produce al multiplicar la fila 3 de A por 5, entonces $\det B = 5 \cdot \det A$.

h. Si B se forma al sumar a una fila de A una combinación lineal de las otras filas, entonces $\det B = \det A$.

i. $\det A^T = -\det A$.

j. $\det(-A) = -\det A$.

k. $\det A^TA \geq 0$.

l. Cualquier sistema de n ecuaciones lineales con n variables puede resolverse mediante la regla de Cramer.

m. Si $\mathbf{u}$ y $\mathbf{v}$ están en $\mathbb{R}^2$ y $\det [\mathbf{u} \quad \mathbf{v}] = 10$, entonces el área del triángulo en el plano con vértices en $\mathbf{0}$, $\mathbf{u}$ y $\mathbf{v}$ es 10.

n. Si $A^3 = 0$, entonces $\det A = 0$.

o. Si A es invertible, entonces $\det A^{-1} = \det A$.

p. Si A es invertible, entonces $(\det A)(\det A^{-1}) = 1$.

Utilice operaciones por fila para mostrar que todos los determinantes de los siguientes ejercicios 2, 3 y 4 son cero.

2. $\begin{vmatrix} 12 & 13 & 14 \\ 15 & 16 & 17 \\ 18 & 19 & 20 \end{vmatrix}$

3. $\begin{vmatrix} 1 & a & b+c \\ 1 & b & a+c \\ 1 & c & a+b \end{vmatrix}$

4. $\begin{vmatrix} a & b & c \\ a+x & b+x & c+x \\ a+y & b+y & c+y \end{vmatrix}$

Encuentre los determinantes de los ejercicios 5 y 6.

5. $\begin{vmatrix} 9 & 1 & 9 & 9 & 9 \\ 9 & 0 & 9 & 9 & 2 \\ 4 & 0 & 0 & 5 & 0 \\ 9 & 0 & 3 & 9 & 0 \\ 6 & 0 & 0 & 7 & 0 \end{vmatrix}$

6. $\begin{vmatrix} 4 & 8 & 8 & 8 & 5 \\ 0 & 1 & 0 & 0 & 0 \\ 6 & 8 & 8 & 8 & 7 \\ 0 & 8 & 8 & 3 & 0 \\ 0 & 8 & 2 & 0 & 0 \end{vmatrix}$

7. Muestre que la ecuación de la línea en $\mathbb{R}^2$ a través de los distintos puntos (x_1, y_1) y (x_2, y_2) puede escribirse como

$$\det \begin{bmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{bmatrix} = 0.$$

8. Encuentre una ecuación de determinantes 3×3 similar a la del ejercicio 7 para describir la ecuación de la línea que pasa por (x_1, y_1) con pendiente m.

Los ejercicios 9 y 10 se refieren a los determinantes de las siguientes *matrices de Vandermonde.*

$$T = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}, \qquad V(t) = \begin{bmatrix} 1 & t & t^2 & t^3 \\ 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \end{bmatrix}$$

9. Utilice operaciones por fila para mostrar que

$\det T = (b - a)(c - a)(c - b)$.

10. Sea $f(t) = \det V$, con x_1, x_2 y x_3 distintos entre sí. Explique por qué $f(t)$ es un polinomio cúbico, muestre que el coeficiente de t^3 es diferente de cero, y encuentre tres puntos sobre la gráfica de f.

11. Calcule el área del paralelogramo determinado por los puntos (1, 4), (−1, 5), (3, 9), y (5, 8). ¿Cómo puede afirmar usted que el cuadrilátero determinado por los puntos es realmente un paralelogramo?

12. Utilice el concepto de área de un paralelogramo para escribir un enunciado acerca de una matriz A de 2×2 que sea cierto si, y sólo si, A es invertible.

13. Muestre que si A es invertible, entonces adj A es invertible y

$$(\text{adj } A)^{-1} = \frac{1}{\det A} A .$$

[*Sugerencia:* Dadas las matrices B y C, ¿qué cálculo(s) mostraría(n) que C es el inverso de B?]

14. Sean A, B, C, D e I matrices de $n \times n$. Utilice la definición o las propiedades de un determinante para justificar las siguientes fórmulas. El inciso (c) es útil en las aplicaciones de valores propios (capítulo 5).

a. $\det \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} = \det A$ b. $\det \begin{bmatrix} I & 0 \\ C & D \end{bmatrix} = \det D$

c. $\det \begin{bmatrix} A & 0 \\ C & D \end{bmatrix} = (\det A)(\det D) = \det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$

15. Sean A, B, C y D matrices de $n \times n$ con A invertible.

a. Encuentre las matrices X e Y para producir la factorización LU en bloques

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} A & B \\ 0 & Y \end{bmatrix}$$

y luego muestre que

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = (\det A) \cdot \det(D - CA^{-1}B).$$

b. Muestre que si $AC = CA$, entonces

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - CB) .$$

16. Sea J la matriz de $n \times n$ con sólo números uno, y considere $A = (a - b)I + bJ$; esto es,

$$A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Confirme que $\det A = (a - b)^{n-1}[a + (n - 1)b]$ de la siguiente manera:

a. Reste la fila 2 a la fila 1, la fila 3 a la fila 2, y así sucesivamente, y explique por qué esto no cambia el determinante de la matriz.

b. Con la matriz resultante de (a), sume la columna 1 a la columna 2, después sume esta nueva columna 2 a la columna 3, y así sucesivamente, y explique por qué esto no cambia el determinante.

c. Encuentre el determinante de la matriz que resultó en (b).

17. Sea A la matriz original dada en el ejercicio 16, y sean

$$B = \begin{bmatrix} a - b & b & b & \cdots & b \\ 0 & a & b & \cdots & b \\ 0 & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b & b & \cdots & a \end{bmatrix},$$

$$C = \begin{bmatrix} b & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Observe que A, B y C son casi iguales excepto por que la primera columna de A es igual a la suma de las primeras columnas de B y C. Una *propiedad de linealidad* de la función

determinante, analizada en la sección 3.2, establece que det $A = \det B + \det C$. Use este hecho para probar por inducción la fórmula del ejercicio 16 sobre el tamaño de la matriz A.

18. **[M]** Aplique el resultado del ejercicio 16 para encontrar los determinantes de las matrices siguientes, y confirme sus respuestas mediante el uso de un programa de matrices.

$$\begin{bmatrix} 3 & 8 & 8 & 8 \\ 8 & 3 & 8 & 8 \\ 8 & 8 & 3 & 8 \\ 8 & 8 & 8 & 3 \end{bmatrix} \quad \begin{bmatrix} 8 & 3 & 3 & 3 & 3 \\ 3 & 8 & 3 & 3 & 3 \\ 3 & 3 & 8 & 3 & 3 \\ 3 & 3 & 3 & 8 & 3 \\ 3 & 3 & 3 & 3 & 8 \end{bmatrix}$$

19. **[M]** Use un programa de matrices para calcular los determinantes de las siguientes matrices.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

Use los resultados para estimar el determinante de la siguiente matriz, y confirme su estimación utilizando operaciones por fila para evaluar ese determinante.

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}$$

20. **[M]** Aplique el método del ejercicio 19 para estimar el determinante de

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 3 & 3 & \cdots & 3 \\ 1 & 3 & 6 & \cdots & 6 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 3 & 6 & \cdots & 3(n-1) \end{bmatrix}$$

Justifique su conjetura. [*Sugerencia:* Use el ejercicio 14(c) y el resultado del ejercicio 19.]

4

Espacios vectoriales

EJEMPLO INTRODUCTORIO

Vuelo espacial y sistemas de control

Con doce pisos de altura y peso de 75 toneladas, el *Columbia* se elevó majestuosamente desde la plataforma de lanzamiento en una fresca mañana de abril de 1981. El primer transbordador de Estados Unidos, producto de diez años de investigación, fue un triunfo de la ingeniería de sistemas de control que abarca muchas ramas ingenieriles —aeronáutica, química, eléctrica, hidráulica y mecánica.

Los sistemas de control del transbordador espacial resultan absolutamente críticos para el vuelo. Como el transbordador tiene un fuselaje inestable, requiere de constante vigilancia por computadora durante el vuelo atmosférico. Los sistemas de control de vuelo envían una corriente de comandos a las superficies de control aerodinámicas y a 44 pequeños impulsores de propulsión a chorro. En la figura 1 se muestra un típico sistema con retroalimentación en ciclo cerrado que controla el ángulo de inclinación de la punta de la nariz del transbordador durante el vuelo. Los símbolos de empalme ($\otimes$) muestran dónde se añaden las señales de diversos sensores a las señales de la computadora que fluyen por la parte superior de la figura.

Matemáticamente, las señales de entrada y salida de un sistema de control son funciones. Es importante, para las aplicaciones, que estas señales puedan sumarse, como en la figura 1, y multiplicarse por escalares. Estas dos operaciones con funciones tienen propiedades algebraicas completamente análogas a las operaciones de suma de vectores en $\mathbb{R}^n$ y multiplicación de un vector por un escalar, como se verá en las secciones 4.1 y 4.8. Por esta razón, al conjunto de todas las posibles entradas (funciones) se le denomina *espacio vectorial.* Los fundamentos matemáticos de la ingeniería de sistemas descansan sobre los espacios vectoriales y las funciones, y en este capítulo se amplía la teoría de vectores en $\mathbb{R}^n$ para incluir tales funciones. Después, se verá cómo surgen otros espacios vectoriales en ingeniería, física y estadística.

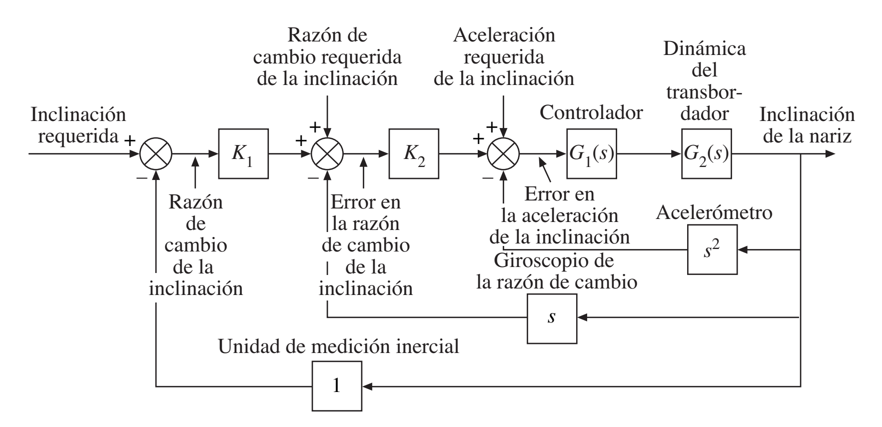


FIGURA 1 Sistemas de control para el transbordador espacial. (*Fuente: Control Systems Engineering*, por Norman S. Nise, Benjamin-Cummings Publishing, 1992, pág. 274. Esquema simplificado basado en *Space Shuttle GN&C Operations Manual*, Rockwell International, 1988.)

Las semillas matemáticas sembradas en los capítulos 1 y 2 germinarán y comenzarán a florecer en este capítulo. La belleza y el poder del álgebra se verán con mayor claridad cuando perciba a $\mathbb{R}^n$ como sólo uno de los diversos espacios vectoriales que surgen de manera natural en problemas de aplicación. En realidad, el estudio de los espacios vectoriales no es demasiado diferente del propio estudio de $\mathbb{R}^n$, porque es posible usar la experiencia geométrica adquirida con $\mathbb{R}^2$ y $\mathbb{R}^3$ para visualizar muchos conceptos generales.

En este capítulo se iniciará con las definiciones básicas de la sección 4.1, para después desarrollar gradualmente el marco general de los espacios vectoriales. Una meta de las secciones 4.3, 4.4 y 4.5 es mostrar lo mucho que otros espacios vectoriales se parecen a $\mathbb{R}^n$. La sección 4.6, que trata acerca del rango, es uno de los puntos principales del capítulo, ahí se usa terminología de espacios vectoriales para vincular importantes hechos acerca de las matrices rectangulares. En la sección 4.8 se aplicará la teoría del capítulo a las señales discretas y a las ecuaciones en diferencias que se usan en sistemas de control digitales como los del transbordador espacial. Las cadenas de Markov, en la sección 4.9, ofrecerán un cambio de paso con respecto a las secciones más teóricas del capítulo, y proporcionarán buenos ejemplos para los conceptos que se introducirán en el capítulo 5.

4.1 ESPACIOS Y SUBESPACIOS VECTORIALES

Gran parte de la teoría presentada en los capítulos 1 y 2 se basó en ciertas propiedades algebraicas simples y evidentes de $\mathbb{R}^n$, las cuales se enlistaron en la sección 1.3. De hecho, muchos otros sistemas matemáticos poseen las mismas propiedades. Las propiedades específicas de interés se enlistan en la siguiente definición.

DEFINICIÓN

Un **espacio vectorial** es un conjunto no vacío V de objetos, llamados *vectores*, en el que están definidas dos operaciones, llamadas *suma* y *multiplicación por escalares* (números reales), sujetas a los diez axiomas (o reglas) que se enlistan a continuación.[1] Los axiomas deben ser válidos para todos los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ en V y todos los escalares c y d.

1. La suma de $\mathbf{u}$ y $\mathbf{v}$, denotada mediante $\mathbf{u} + \mathbf{v}$, está en V.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. Existe un vector **cero** $\mathbf{0}$ en V tal que $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. Para cada $\mathbf{u}$ en V, existe un vector $-\mathbf{u}$ en V tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. El múltiplo escalar de $\mathbf{u}$ por c, denotado mediante $c\mathbf{u}$, está en V.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

Mediante estos axiomas, es posible demostrar que el vector cero del axioma 4 es único, y que el vector $-\mathbf{u}$, llamado el **negativo** de $\mathbf{u}$, del axioma 5 es único para cada $\mathbf{u}$ en V. Vea los ejercicios 25 y 26. En los ejercicios de este capítulo también se delinearán demostraciones de los siguientes hechos sencillos:

Para cada $\mathbf{u}$ en V y escalar c,

$$0\mathbf{u} = \mathbf{0} \quad (1)$$

$$c\mathbf{0} = \mathbf{0} \quad (2)$$

$$-\mathbf{u} = (-1)\mathbf{u} \quad (3)$$

EJEMPLO 1 Los espacios $\mathbb{R}^n$, donde $n \geq 1$, son los principales ejemplos de espacios vectoriales. La intuición geométrica desarrollada para $\mathbb{R}^3$ resultará muy útil para entender y visualizar muchos conceptos a lo largo del capítulo.

EJEMPLO 2 Sea V el conjunto de todas las flechas (segmentos de líneas dirigidos) presentes en el espacio tridimensional; dos de estas flechas se consideran iguales si tienen la misma longitud y apuntan en la misma dirección. La suma se define por medio de la regla del paralelogramo (de la sección 1.3), y para cada $\mathbf{v}$ en V se define $c\mathbf{v}$ como la flecha cuya longitud es $|c|$ veces la longitud de $\mathbf{v}$, y que apunta en la misma dirección que $\mathbf{v}$ si $c \geq 0$ y en la dirección opuesta en caso contrario. (Vea la figura 1.) Muestre que V es un espacio vectorial. Este espacio es un modelo común en problemas de física para diversas fuerzas.

v 3v −v

FIGURA 1

[1]Técnicamente, V es un *espacio vectorial real.* Toda la teoría de este capítulo es válida también para *espacios vectoriales complejos*, donde los escalares son números complejos. Esto se estudiará de manera breve en el capítulo 5. Hasta entonces, se supondrá que todos los escalares son reales.

Solución La definición de V es geométrica, y utiliza conceptos de longitud y dirección. No intervienen coordenadas xyz. Una flecha de longitud cero es un solo punto y representa el vector cero. El negativo de $\mathbf{v}$ es $(-1)\mathbf{v}$. Así, los axiomas 1, 4, 5, 6 y 10 son evidentes; los demás se verifican geométricamente. Por ejemplo, vea las figuras 2 y 3.

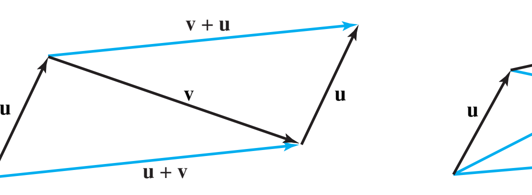


FIGURA 2 $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

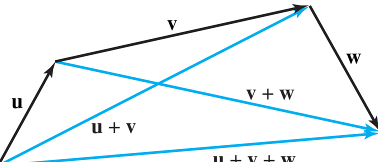


FIGURA 3 $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

EJEMPLO 3 Sea $\mathbb{S}$ el espacio de todas las sucesiones infinitas de números a derecha e izquierda (normalmente escritas en fila y no en columna):

$$\{y_k\} = (\ldots, y_{-2}, y_{-1}, y_0, y_1, y_2, \ldots)$$

Si $\{z_k\}$ es otro elemento de $\mathbb{S}$, entonces la suma $\{y_k\} + \{z_k\}$ es la sucesión $\{y_k + z_k\}$ que se forma al sumar términos correspondientes de $\{y_k\}$ y $(z_k\}$. El múltiplo escalar $c\{y_k\}$ es la sucesión $\{cy_k\}$. Los axiomas de espacio vectorial se verifican de la misma forma que se hizo para $\mathbb{R}^n$.

Los elementos de $\mathbb{S}$ aparecen en ingeniería, por ejemplo, siempre que una señal se mide (o muestrea) en tiempos discretos. Una señal puede ser eléctrica, mecánica, óptica, etc. Los sistemas de control principales del transbordador espacial, mencionados en la introducción del capítulo, usan señales discretas (o digitales). Por con conveniencia, se llamará a $\mathbb{S}$ espacio de **señales** (de tiempo discreto). Una señal puede visualizarse por medio de una gráfica como la de la figura 4.

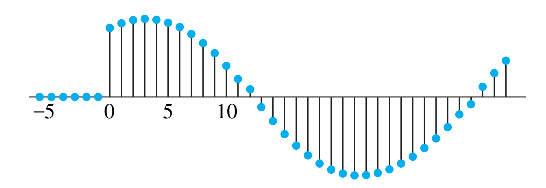


FIGURA 4 Una señal de tiempo discreto.

EJEMPLO 4 Para $n \geq 0$, el conjunto $\mathbb{P}_n$ de polinomios de grado n o menor consiste en todos los polinomios de la forma

$$\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \tag{4}$$

donde los coeficientes $a_0, \ldots, a_n$ y la variable t son números reales. El *grado* de $\mathbf{p}$ es la mayor potencia de t en (4) cuyo coeficiente no es cero. Si $\mathbf{p}(t) = a_0 \neq 0$, el grado de $\mathbf{p}$

es cero. Si todos los coeficientes son cero, $\mathbf{p}$ es el *polinomio cero*. El polinomio cero está incluido en $\mathbb{P}_n$ aun cuando su grado, por razones técnicas, no esté definido.

Si $\mathbf{p}$ está dado por (4), y si $\mathbf{q}(t) = b_0 + b_1 t + \cdots + b_n t^n$, entonces la suma $\mathbf{p} + \mathbf{q}$ se define mediante

$$\begin{aligned}(\mathbf{p}+\mathbf{q})(t) &= \mathbf{p}(t) + \mathbf{q}(t) \\ &= (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n\end{aligned}$$

El múltiplo escalar $c\mathbf{p}$ es el polinomio definido por

$$(c\mathbf{p})(t) = c\mathbf{p}(t) = ca_0 + (ca_1)t + \cdots + (ca_n)t^n$$

Estas definiciones satisfacen los axiomas 1 y 6 porque $\mathbf{p} + \mathbf{q}$ y $c\mathbf{p}$ son polinomios de grado menor o igual que n. Los axiomas 2, 3, y del 7 al 10 se siguen a partir de las propiedades de los números reales. Resulta claro que el polinomio cero actúa como el vector cero del axioma 4. Por último, $(-1)\mathbf{p}$ actúa como el negativo de $\mathbf{p}$, y se cumple el axioma 5. Por lo tanto, $\mathbb{P}_n$ es un espacio vectorial.

Los espacios vectoriales $\mathbb{P}_n$ para diferentes n se usan, por ejemplo, en el análisis de tendencia estadística de datos, el cual se estudia en la sección 6.8.

EJEMPLO 5 Sea V el conjunto de todas las funciones que dan valores reales definidas para un conjunto $\mathbb{D}$. (Por lo general, $\mathbb{D}$ se toma como el conjunto de los números reales o como algún intervalo de la recta real.) Las funciones se suman de la forma acostumbrada: $\mathbf{f} + \mathbf{g}$ es la función cuyo valor en t en el dominio $\mathbb{D}$ es $\mathbf{f}(t) + \mathbf{g}(t)$. De igual manera, para un escalar c y una $\mathbf{f}$ en V, el múltiplo escalar $c\mathbf{f}$ es la función cuyo valor en t es $c\mathbf{f}(t)$. Por ejemplo, si $\mathbb{D} = \mathbb{R}$, $\mathbf{f}(t) = 1 + \text{sen}\, 2t$, y $\mathbf{g}(t) = 2 + .5t$, entonces

$$(\mathbf{f}+\mathbf{g})(t) = 3 + \text{sen}\, 2t + .5t \quad \text{y} \quad (2\mathbf{g})(t) = 4 + t$$

Dos funciones en V son iguales si, y sólo si, sus valores coinciden para cada t en $\mathbb{D}$. Por lo tanto, el vector cero en V es la función que es idénticamente cero, $\mathbf{f}(t) = 0$ para toda t, y el negativo de $\mathbf{f}$ es $(-1)\mathbf{f}$. Los axiomas 1 y 6 son evidentemente ciertos, y los otros axiomas se siguen a partir de las propiedades de los números reales, así que V es un espacio vectorial.

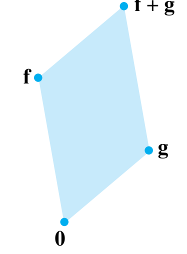


FIGURA 5 La suma de dos vectores (funciones).

Es importante pensar en cada función en el espacio vectorial V del ejemplo 5 como un solo objeto, un solo "punto" o vector en el espacio vectorial. La suma de dos vectores $\mathbf{f}$ y $\mathbf{g}$ (funciones en V o elementos de *cualquier* espacio vectorial) puede visualizarse como en la figura 5, porque esto puede ayudar a aplicar a un espacio vectorial general la intuición geométrica adquirida al trabajar con el espacio vectorial $\mathbb{R}^n$. Como ayuda, puede consultar la *guía de estudio* (*Study Guide*) mientras aprende a adoptar este punto de vista más general.

Subespacios

En muchos problemas, un espacio vectorial consta de un subconjunto adecuado de vectores de algún espacio vectorial mayor. En este caso, será necesario verificar sólo tres de los diez axiomas de espacios vectoriales. El resto quedarán satisfechos de manera automática.

DEFINICIÓN

Un **subespacio** de un espacio vectorial V es un subconjunto H de V que tiene tres propiedades:

a. El vector cero de V está en H.[2]

b. H es cerrado bajo la suma de vectores. Esto es, para cada $\mathbf{u}$ y $\mathbf{v}$ en H, la suma $\mathbf{u} + \mathbf{v}$ está en H.

c. H es cerrado bajo la multiplicación por escalares. Esto es, para cada $\mathbf{u}$ en H y cada escalar c, el vector $c\mathbf{u}$ está en H.

Las propiedades (a), (b) y (c) garantizan que un subespacio H de V es en sí mismo un *espacio vectorial*, bajo las operaciones de espacio vectorial ya definidas en V. Para verificar esto, observe que las propiedades (a), (b) y (c) son los axiomas 1, 4 y 6. Los axiomas 2, 3, y del 7 al 10 son verdaderos de manera automática en H porque se aplican a todos los elementos de V, incluidos aquellos que están en H. El axioma 5 también es verdadero en H, porque si $\mathbf{u}$ está en H, entonces $(-1)\mathbf{u}$ está en H según (c), y por la ecuación (3) de la página 217 se sabe que $(-1)\mathbf{u}$ es el vector $-\mathbf{u}$ del axioma 5.

Así, todo subespacio es un espacio vectorial. De manera recíproca, todo espacio vectorial es un subespacio (de sí mismo o posiblemente de espacios mayores). El término *subespacio* es usado cuando se consideran por lo menos dos espacios, con uno dentro de otro, y la frase *subespacio de V* identifica a V como el espacio más grande. (Vea la figura 6.)

FIGURA 6 Un subespacio de V.

EJEMPLO 6 El conjunto que consta de únicamente el vector cero en un espacio vectorial V es un subespacio de V, llamado **subespacio cero** y que se escribe como $\{\mathbf{0}\}$.

EJEMPLO 7 Sea $\mathbb{P}$ el conjunto de todos los polinomios con coeficientes reales, con operaciones en $\mathbb{P}$ definidas igual que para las funciones. Entonces $\mathbb{P}$ es un subespacio del espacio de todas las funciones que producen un valor real definidas en $\mathbb{R}$. También, para cada $n \geq 0$, $\mathbb{P}_n$ es un subespacio de $\mathbb{P}$, porque $\mathbb{P}_n$ es un subconjunto de $\mathbb{P}$ que contiene al polinomio cero, la suma de dos polinomios en $\mathbb{P}_n$ también está en $\mathbb{P}_n$, y un múltiplo escalar de un polinomio en $\mathbb{P}_n$ también está en $\mathbb{P}_n$.

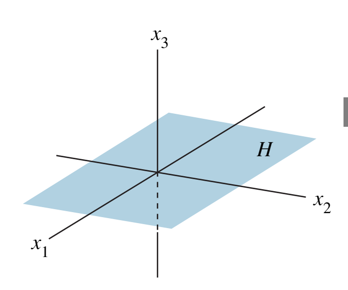


FIGURA 7 Un plano x_1x_2 es un subespacio de $\mathbb{R}^3$.

EJEMPLO 8 El espacio vectorial $\mathbb{R}^2$ *no* es un subespacio de $\mathbb{R}^3$ porque $\mathbb{R}^2$ ni siquiera es un subconjunto de $\mathbb{R}^3$. (Todos los vectores en $\mathbb{R}^3$ tienen tres entradas, mientras que los vectores en $\mathbb{R}^2$ tienen sólo dos.) El conjunto

$$H = \left\{ \begin{bmatrix} s \\ t \\ 0 \end{bmatrix} : s \text{ y } t \text{ son reales} \right\}$$

es un subconjunto de $\mathbb{R}^3$ que "se ve" y "actúa" como $\mathbb{R}^2$, aunque es lógicamente distinto de $\mathbb{R}^2$. Vea la figura 7. Demuestre que H es un subespacio de $\mathbb{R}^3$.

Solución El vector cero está en H, y H es cerrado bajo la suma de vectores y la multiplicación por escalares porque estas operaciones con vectores en H producen siempre vectores cuyas terceras entradas son cero (y por ende pertenecen a H). Entonces H es un subespacio de $\mathbb{R}^3$.

[2] Algunos textos reemplazan la propiedad (a) de esta definición por el supuesto de que H no es vacío. Entonces (a) podría deducirse de (c) y de que $0\mathbf{u} = \mathbf{0}$. Pero la mejor manera de comprobar para un subespacio es buscar inicialmente el vector cero. Si $\mathbf{0}$ está en H, entonces deben verificarse las propiedades (b) y (c). Si $\mathbf{0}$ *no* está en H, entonces H no puede ser un subespacio, y ya no hace falta confirmar las otras propiedades.

EJEMPLO 9 Un plano en $\mathbb{R}^3$ que *no* pasa por el origen no es un subespacio de $\mathbb{R}^3$, porque el plano no contiene al vector cero de $\mathbb{R}^3$. De manera similar, una línea en $\mathbb{R}^2$ que *no* pasa por el origen, como en la figura 8, *no* es un subespacio de $\mathbb{R}^2$.

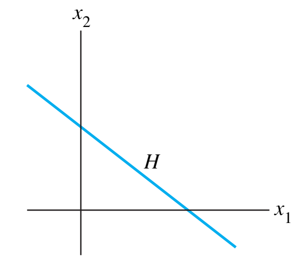


FIGURA 8
Una línea que no es un espacio vectorial.

Un subespacio generado por un conjunto

El ejemplo siguiente ilustra una de las maneras más comunes de describir un subespacio. Al igual que en el capítulo 1, el término **combinación lineal** se refiere a cualquier suma de múltiplos escalares de vectores, y Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ denota el conjunto de todos los vectores que pueden escribirse como combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

EJEMPLO 10 Dados $\mathbf{v}_1$ y $\mathbf{v}_2$ en un espacio vectorial V, sea $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$. Demuestre que H es un subespacio de V.

Solución El vector cero está en H, dado que $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$. Para mostrar que H es cerrado bajo la suma de vectores, tome dos vectores arbitrarios en H. por ejemplo,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 \quad \text{y} \quad \mathbf{w} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

De acuerdo con los axiomas 2, 3 y 8 para el espacio vectorial V,

$$\begin{aligned}\mathbf{u} + \mathbf{w} &= (s_1\mathbf{v}_1 + s_2\mathbf{v}_2) + (t_1\mathbf{v}_1 + t_2\mathbf{v}_2) \\ &= (s_1 + t_1)\mathbf{v}_1 + (s_2 + t_2)\mathbf{v}_2\end{aligned}$$

Entonces $\mathbf{u} + \mathbf{w}$ está en H. Además, si c es un escalar, entonces, por los axiomas 7 y 9,

$$c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$$

lo cual muestra que $c\mathbf{u}$ está en H, y H es cerrado bajo la multiplicación por escalares. Entonces H es un subespacio de V.

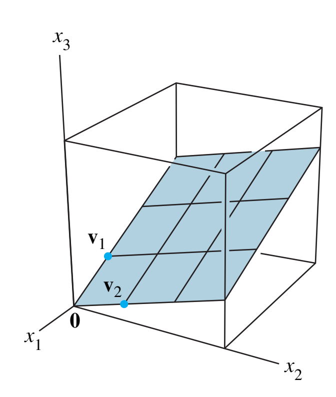
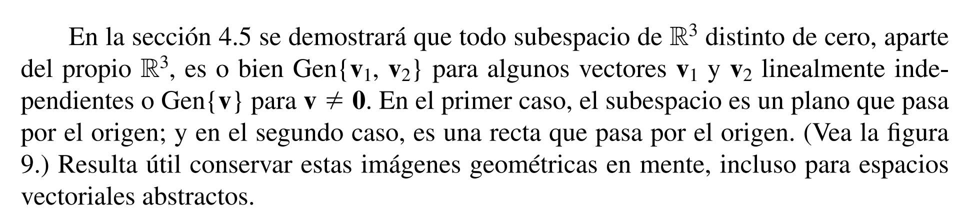


FIGURA 9
Ejemplo de un subespacio.



En la sección 4.5 se demostrará que todo subespacio de $\mathbb{R}^3$ distinto de cero, aparte del propio $\mathbb{R}^3$, es o bien Gen$\{\mathbf{v}_1, \mathbf{v}_2\}$ para algunos vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ linealmente independientes o Gen$\{\mathbf{v}\}$ para $\mathbf{v} \neq \mathbf{0}$. En el primer caso, el subespacio es un plano que pasa por el origen; y en el segundo caso, es una recta que pasa por el origen. (Vea la figura 9.) Resulta útil conservar estas imágenes geométricas en mente, incluso para espacios vectoriales abstractos.

El argumento del ejemplo 10 se generaliza fácilmente para demostrar el teorema siguiente.

TEOREMA 1 Si $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en un espacio vectorial V, entonces Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es un subespacio de V.

A Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ se le llama el **subespacio generado** por $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$. Dado cualquier subespacio H de V, un **conjunto generador** para H es un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en H tal que $H = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

El ejemplo siguiente muestra cómo usar el teorema 1.

EJEMPLO 11 Sea H el conjunto de todos los vectores de la forma $(a - 3b, b - a, a, b)$, donde a y b son escalares arbitrarios. Esto es, sea $H = \{(a - 3b, b - a, a, b)$: a y b en $\mathbb{R}\}$. Demuestre que H es un subespacio de $\mathbb{R}^4$.

Solución Escriba los vectores de H como vectores columna. Entonces un vector arbitrario en H tiene la forma

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a\underset{\uparrow \atop \mathbf{v}_1}{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}} + b\underset{\uparrow \atop \mathbf{v}_2}{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}}$$

Este cálculo muestra que $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_1$ y $\mathbf{v}_2$ son los vectores indicados anteriormente. Entonces H es un subespacio de $\mathbb{R}^4$ de acuerdo con el teorema 1.

En el ejemplo 11 se ilustra una técnica útil para expresar un subespacio H como el conjunto de combinaciones lineales de alguna pequeña colección de vectores. Si $H = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$, se puede pensar en los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_p$ del conjunto generador como "asas" que permiten manipular el subespacio H. A menudo, cálculos con la infinidad de elementos de H se reduce a operaciones con el número finito de vectores del conjunto generador.

EJEMPLO 12 Encuentre el o los valores de h para los cuales $\mathbf{y}$ está en el subespacio de $\mathbb{R}^3$ generado por $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, si

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

Solución Esta pregunta corresponde al problema de práctica 2 de la sección 1.3, escrito aquí usando el término *subespacio* en lugar de $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. La solución obtenida anteriormente muestra que $\mathbf{y}$ está en $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si, y sólo si, $h = 5$. Ahora es recomendable repasar esa solución, y también los ejercicios del 11 al 14 y del 17 al 21 de la sección 1.3.

Aunque muchos espacios vectoriales de este capítulo son subespacios de $\mathbb{R}^n$, es importante recordar que la teoría abstracta se puede aplicar también a otros espacios vectoriales. Los espacios vectoriales de funciones surgen en muchas aplicaciones, y posteriormente se les prestará más atención.

PROBLEMAS DE PRÁCTICA

1. Muestre que el conjunto H de los puntos en $\mathbb{R}^2$ de la forma $(3s, 2 + 5s)$ no es un espacio vectorial, mostrando que no es cerrado bajo la multiplicación por escalares. (Encuentre un vector específico $\mathbf{u}$ en H, y un escalar c tal que $c\mathbf{u}$ no esté en H.)
2. Sea $W = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$, donde $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en un espacio vectorial V. Muestre que $\mathbf{v}_k$ está en W para $1 \leq k \leq p$. [*Sugerencia:* Primero escriba una ecuación donde muestre que $\mathbf{v}_1$ está en W. Luego ajuste la notación para el caso general.]

4.1 EJERCICIOS

1. Sea V el primer cuadrante en el plano xy; esto es, sea

$$V = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x \geq 0,\, y \geq 0 \right\}$$

 a. Si $\mathbf{u}$ y $\mathbf{v}$ están en V, ¿está $\mathbf{u} + \mathbf{v}$ en V? ¿Por qué?

 b. Encuentre un vector específico $\mathbf{u}$ en V y un escalar específico tal que $c\mathbf{u}$ *no* esté en V. (Esto basta para demostrar que V *no* es un espacio vectorial.)

2. Sea W la unión del primer y tercer cuadrantes en el plano xy. Esto es, sea $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : xy \geq 0 \right\}$.

 a. Si $\mathbf{u}$ está en W y c es cualquier escalar, ¿está $c\mathbf{u}$ en W? ¿Por qué?

 b. Encuentre vectores específicos $\mathbf{u}$ y $\mathbf{v}$ en W tales que $\mathbf{u} + \mathbf{v}$ no esté en W. Esto basta para demostrar que W *no* es un espacio vectorial.

3. Sea H el conjunto de puntos que están dentro del círculo unitario en el plano xy. Esto es, sea $H = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x^2 + y^2 \leq 1 \right\}$.

 Encuentre un ejemplo específico —dos vectores o un vector y un escalar— para mostrar que H no es un subespacio de $\mathbb{R}^2$.

4. Construya una figura geométrica para ilustrar por qué una línea en $\mathbb{R}^2$ que *no* pasa por el origen no es cerrada bajo la suma de vectores.

En los ejercicios 5 a 8, determine si el conjunto dado es un subespacio de $\mathbb{P}_n$ para algún valor adecuado de n. Justifique sus respuestas.

5. Todos los polinomios de la forma $\mathbf{p}(t) = at^2$, donde a está en $\mathbb{R}$.

6. Todos los polinomios de la forma $\mathbf{p}(t) = a + t^2$, donde a está en $\mathbb{R}$.

7. Todos los polinomios de grado 3 o menor, con coeficientes enteros.

8. Todos los polinomios en $\mathbb{P}_n$, tales que $\mathbf{p}(0) = 0$.

9. Sea H el conjunto de todos los vectores de la forma $\begin{bmatrix} s \\ 3s \\ 2s \end{bmatrix}$.

 Encuentre un vector $\mathbf{v}$ en $\mathbb{R}^3$ tal que $H = \text{Gen}\{\mathbf{v}\}$. ¿Por qué muestra esto que H es un subespacio de $\mathbb{R}^3$?

10. Sea H el conjunto de todos los vectores de la forma $\begin{bmatrix} 2t \\ 0 \\ -t \end{bmatrix}$.

 Muestre que H es un subespacio de $\mathbb{R}^3$. (Use el método del ejercicio 9.)

11. Sea W el conjunto de todos los vectores de la forma $\begin{bmatrix} 5b + 2c \\ b \\ c \end{bmatrix}$, donde b y c son arbitrarios. Encuentre vectores $\mathbf{u}$ y $\mathbf{v}$ tales que $W = \text{Gen}\{\mathbf{u}, \mathbf{v}\}$. ¿Por qué muestra esto que W es un subespacio de $\mathbb{R}^3$?

12. Sea W el conjunto de todos los vectores de la forma $\begin{bmatrix} s + 3t \\ s - t \\ 2s - t \\ 4t \end{bmatrix}$.

 Muestre que W es un subespacio de $\mathbb{R}^4$. (Use el método del ejercicio 11.)

13. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}$, y $\mathbf{w} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

 a. ¿Está $\mathbf{w}$ en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Cuántos vectores hay en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

 b. ¿Cuántos vectores hay en $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

 c. ¿Está $\mathbf{w}$ en el subespacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Por qué?

14. Sean $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ como en el ejercicio 13, y sea $\mathbf{w} = \begin{bmatrix} 8 \\ 4 \\ 7 \end{bmatrix}$. ¿Está $\mathbf{w}$ en el subespacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Por qué?

En los ejercicios 15 a 18, sea W el conjunto de todos los vectores de la forma que se muestra, donde a, b y c representan números reales arbitrarios. En cada caso, encuentre un conjunto S de vectores que genere W o proporcione un ejemplo para demostrar que W *no* es un espacio vectorial.

15. $\begin{bmatrix} 3a + b \\ 4 \\ a - 5b \end{bmatrix}$

16. $\begin{bmatrix} -a + 1 \\ a - 6b \\ 2b + a \end{bmatrix}$

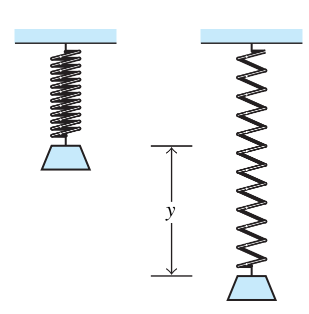
17. $\begin{bmatrix} a - b \\ b - c \\ c - a \\ b \end{bmatrix}$

18. $\begin{bmatrix} 4a + 3b \\ 0 \\ a + b + c \\ c - 2a \end{bmatrix}$

19. Si una masa m se coloca en el extremo de un resorte y se jala de ella hacia abajo y luego se le suelta, el sistema de masa-resorte comenzará a oscilar. El desplazamiento y de la masa desde su posición de reposo está dado por una función de la forma

$$y(t) = c_1 \cos \omega t + c_2 \operatorname{sen} \omega t \tag{5}$$

 donde ω es una constante que depende del resorte y de la masa. Demuestre que el conjunto de todas las funciones descritas en (5) (con ω fija y c_1 y c_2 arbitrarias) es un espacio vectorial.



20. El conjunto de todas las funciones continuas con valores reales, definidas en un intervalo cerrado $[a, b]$ en $\mathbb{R}$, se denota mediante $C[a, b]$. Este conjunto es un subespacio del espacio vectorial de todas las funciones con valores reales, definidas en $[a, b]$.

a. ¿Qué hechos acerca de las funciones continuas deben verificarse para demostrar que $C[a, b]$ es en realidad un subespacio vectorial como se asegura? (Por lo general, estos hechos se estudian en una clase de cálculo.)

b. Demuestre que $\{\mathbf{f} \text{ en } C[a, b] : \mathbf{f}(a) = \mathbf{f}(b)\}$ es un subespacio de $C[a, b]$.

Para enteros positivos fijos m y n, el conjunto $M_{m\times n}$ de todas las matrices de $m \times n$ es un espacio vectorial, bajo las operaciones usuales de suma de matrices y multiplicación por escalares reales.

21. Determine si el conjunto H de todas las matrices de la forma $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ es un subespacio de $M_{2\times 2}$.

22. Sean F una matriz fija de 3×2, y H el conjunto de todas las matrices A en $M_{2\times 4}$ con la propiedad de que $FA = 0$ (la matriz cero en $M_{3\times 4}$). Determine si H es un subespacio de $M_{2\times 4}$.

En los ejercicios 23 y 24 señale cada enunciado como verdadero o falso. Justifique sus respuestas.

23. a. Si $\mathbf{f}$ es una función en el espacio vectorial V de todas las funciones con valores reales definidas en $\mathbb{R}$, y si $\mathbf{f}(t) = 0$ para cualquier t, entonces $\mathbf{f}$ es el vector cero en V.

b. Un vector es una flecha en un espacio tridimensional.

c. Un subconjunto H de un espacio vectorial V es un subespacio de V si el vector cero está en H.

d. Un subespacio también es un espacio vectorial.

e. Se usan señales analógicas en los sistemas de control principales del transbordador espacial, los cuales fueron mencionados en la introducción de este capítulo.

24. a. Un vector es cualquier elemento de un espacio vectorial.

b. Si $\mathbf{u}$ es un vector en un espacio vectorial V, entonces $(-1)\mathbf{u}$ es lo mismo que el negativo de $\mathbf{u}$.

c. Un espacio vectorial es también un subespacio.

d. $\mathbb{R}^2$ es un subespacio de $\mathbb{R}^3$.

e. Un subconjunto H de un espacio vectorial V es un subespacio de V cuando se cumplen las siguientes condiciones: (i) el vector cero de V está en H, (ii) $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} + \mathbf{v}$ están en H, y (iii) c es un escalar y $c\mathbf{u}$ está en H.

Los ejercicios 25 a 29 muestran cómo los axiomas para un espacio vectorial V pueden usarse para demostrar las propiedades elementales que se describieron después de la definición de un espacio vectorial. Llene cada espacio con el número de axioma adecuado. Por el axioma 2, los axiomas 4 y 5 implican, respectivamente, que $\mathbf{0} + \mathbf{u} = \mathbf{u}$ y $-\mathbf{u} + \mathbf{u} = \mathbf{0}$ para toda $\mathbf{u}$.

25. Complete la siguiente demostración de que el vector cero es único. Suponga que $\mathbf{w}$ en V tiene la propiedad de que $\mathbf{u} + \mathbf{w} = \mathbf{w} + \mathbf{u} = \mathbf{u}$ para toda $\mathbf{u}$ en V. En particular, $\mathbf{0} + \mathbf{w} = \mathbf{0}$. Pero $\mathbf{0} + \mathbf{w} = \mathbf{w}$, por el axioma _____. Por lo tanto, $\mathbf{w} = \mathbf{0} + \mathbf{w} = \mathbf{0}$.

26. Complete la siguiente demostración de que $-\mathbf{u}$ es el *único vector* en V tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$. Suponga que $\mathbf{w}$ satisface $\mathbf{u} + \mathbf{w} = \mathbf{0}$. Al sumar $-\mathbf{u}$ en ambos lados, se tiene que

$(-\mathbf{u}) + [\mathbf{u} + \mathbf{w}] = (-\mathbf{u}) + \mathbf{0}$

$[(-\mathbf{u}) + \mathbf{u}] + \mathbf{w} = (-\mathbf{u}) + \mathbf{0}$ por el axioma ______(a)

$\mathbf{0} + \mathbf{w} = (-\mathbf{u}) + \mathbf{0}$ por el axioma ______(b)

$\mathbf{w} = -\mathbf{u}$ por el axioma ______(c)

27. Escriba los números de axioma faltantes en la siguiente demostración de que $0\mathbf{u} = \mathbf{0}$ para cada $\mathbf{u}$ en V.

$0\mathbf{u} = (0 + 0)\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}$ por el axioma _______(a)

Sume el negativo de $0\mathbf{u}$ en ambos lados:

$0\mathbf{u} + (-0\mathbf{u}) = [0\mathbf{u} + 0\mathbf{u}] + (-0\mathbf{u})$

$0\mathbf{u} + (-0\mathbf{u}) = 0\mathbf{u} + [0\mathbf{u} + (-0\mathbf{u})]$ por el axioma _______(b)

$\mathbf{0} = 0\mathbf{u} + \mathbf{0}$ por el axioma _______(c)

$\mathbf{0} = 0\mathbf{u}$ por el axioma _______(d)

28. Escriba los números de axioma faltantes en la siguiente demostración de que $c\mathbf{0} = \mathbf{0}$ para cada escalar c.

$c\mathbf{0} = c(\mathbf{0} + \mathbf{0})$ por el axioma _______(a)

$= c\mathbf{0} + c\mathbf{0}$ por el axioma _______(b)

Sume el negativo de $c\mathbf{0}$ en ambos lados:

$c\mathbf{0} + (-c\mathbf{0}) = [c\mathbf{0} + c\mathbf{0}] + (-c\mathbf{0})$

$c\mathbf{0} + (-c\mathbf{0}) = c\mathbf{0} + [c\mathbf{0} + (-c\mathbf{0})]$ por el axioma _____(c)

$\mathbf{0} = c\mathbf{0} + \mathbf{0}$ por el axioma _____(d)

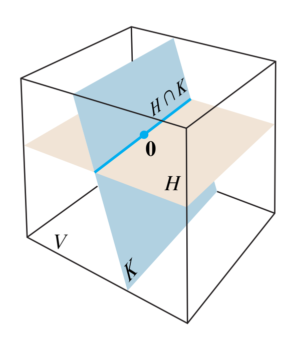
$\mathbf{0} = c\mathbf{0}$ por el axioma _____(e)

29. Demuestre que $(-1)\mathbf{u} = -\mathbf{u}$. [*Sugerencia:* Demuestre que $\mathbf{u} + (-1)\mathbf{u} = \mathbf{0}$. Use algunos axiomas y los resultados de los ejercicios 27 y 26.]

30. Suponga que $c\mathbf{u} = \mathbf{0}$ para algún escalar c distinto de cero. Muestre que $\mathbf{u} = \mathbf{0}$. Mencione los axiomas o propiedades que utilice.

31. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en un espacio vectorial V, y sea H cualquier subespacio de V que contenga tanto a $\mathbf{u}$ como a $\mathbf{v}$. Explique por qué H también contiene a Gen$\{\mathbf{u}, \mathbf{v}\}$. Esto demuestra que Gen$\{\mathbf{u}, \mathbf{v}\}$ es el menor subespacio de V que contiene tanto a $\mathbf{u}$ como a $\mathbf{v}$.

32. Sean H y K subespacios de un espacio vectorial V. La **intersección** de H y K, escrita como $H \cap K$, es el conjunto de los $\mathbf{v}$ en V que pertenece tanto a H como a K. Muestre que $H \cap K$ es un subespacio de V. (Vea la figura.) Dé un ejemplo en $\mathbb{R}^2$ para mostrar que la unión de dos subespacios, en general, no es un subespacio.



33. Dados los subespacios H y K de un espacio vectorial V, la **suma** de H y K, escrita como $H + K$, es el conjunto de todos los vectores en V que puede escribirse como la suma de dos vectores, uno en H y otro en K; esto es,

$H + K = \{\mathbf{w} : \mathbf{w} = \mathbf{u} + \mathbf{v}$ para alguna $\mathbf{u}$ en H y alguna $\mathbf{v}$ en $K\}$

a. Muestre que $H + K$ es un subespacio de V.

b. Muestre que H es un subespacio de $H + K$ y K un subespacio de $H + K$.

34. Suponga que $\mathbf{u}_1, \ldots, \mathbf{u}_p$ y $\mathbf{v}_1, \ldots, \mathbf{v}_q$ son vectores en un espacio vectorial V, y sea

$H = \text{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ y $K = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_q\}$

Muestre que $H + K = \text{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$.

35. **[M]** Muestre que $\mathbf{w}$ está en el subespacio de $\mathbb{R}^4$ generado por $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, donde

$$\mathbf{w} = \begin{bmatrix} -9 \\ 7 \\ 4 \\ 8 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 7 \\ -4 \\ -2 \\ 9 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 5 \\ -1 \\ -7 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 4 \\ 4 \\ -7 \end{bmatrix}$$

36. **[M]** Determine si $\mathbf{y}$ está en el subespacio de $\mathbb{R}^4$ generado por las columnas de A, donde

$$\mathbf{y} = \begin{bmatrix} 6 \\ 7 \\ 1 \\ -4 \end{bmatrix}, \quad A = \begin{bmatrix} 5 & -5 & -9 \\ 8 & 8 & -6 \\ -5 & -9 & 3 \\ 3 & -2 & -7 \end{bmatrix}$$

37. **[M]** El espacio vectorial $H = \text{Gen}\{1, \cos^2 t, \cos^4 t, \cos^6 t\}$ contiene al menos dos funciones interesantes que se usarán en un ejercicio posterior:

$\mathbf{f}(t) = 1 - 8\cos^2 t + 8\cos^4 t$

$\mathbf{g}(t) = -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t$

Estudie la gráfica de $\mathbf{f}$ para $0 \le t \le 2\pi$, y encuentre una fórmula simple para $\mathbf{f}(t)$. Verifique su estimación graficando la diferencia entre $1 + \mathbf{f}(t)$ y su fórmula para $\mathbf{f}(t)$. (Con suerte, se verá la función constante 1.) Repita el procedimiento para $\mathbf{g}$.

38. **[M]** Repita el ejercicio 35 para las funciones

$\mathbf{f}(t) = 3\,\text{sen}\,t - 4\,\text{sen}^3\,t$

$\mathbf{g}(t) = 1 - 8\,\text{sen}^2\,t + 8\,\text{sen}^4\,t$

$\mathbf{h}(t) = 5\,\text{sen}\,t - 20\,\text{sen}^3 t + 16\,\text{sen}^5\,t$

en el espacio vectorial $\text{Gen}\{1, \text{sen}\,t, \text{sen}^2 t, \ldots, \text{sen}^5 t\}$.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Tome cualquier $\mathbf{u}$ en H—por ejemplo, $\mathbf{u} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$—y tome cualquier $c \neq 1$—digamos, $c = 2$. Entonces $c\mathbf{u} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$. Si éste se encuentra en H, entonces hay alguna s tal que

$$\begin{bmatrix} 3s \\ 2 + 5s \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

Esto es, $s = 2$ y $s = 12/5$, lo cual es imposible. Por lo tanto, $2\mathbf{u}$ no está en H, y H no es un espacio vectorial.

2. $\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. Esto expresa $\mathbf{v}_1$ como una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_p$, así que $\mathbf{v}_1$ está en W. En general, $\mathbf{v}_k$ está en W porque

$$\mathbf{v}_k = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{k-1} + 1\mathbf{v}_k + 0\mathbf{v}_{k+1} + \cdots + 0\mathbf{v}_p$$

4.2 ESPACIOS NULOS, ESPACIOS COLUMNA Y TRANSFORMACIONES LINEALES

En aplicaciones de álgebra lineal, los subespacios de $\mathbb{R}^n$ normalmente surgen de dos maneras: (1) como el conjunto de todas las soluciones de un sistema homogéneo de ecuaciones lineales, o (2) como el conjunto de todas las combinaciones lineales de ciertos vectores específicos. En esta sección se compararán y contrastarán las dos descripciones para los subespacios, lo cual permitirá practicar el uso del concepto de subespacio. En realidad, como pronto se descubrirá, ya se ha realizado cierto trabajo con subespacios desde la sección 1.3. La principal característica nueva aquí es la terminología. La sección termina con una explicación del núcleo y del rango de una transformación lineal.

El espacio nulo de una matriz

Considere el siguiente sistema de ecuaciones homogéneas:

$$\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -5x_1 + 9x_2 + x_3 &= 0 \end{aligned} \tag{1}$$

En arreglo matricial, este sistema se escribe como $A\mathbf{x} = \mathbf{0}$, donde

$$A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \tag{2}$$

Recuerde que el conjunto de todas las $\mathbf{x}$ que satisfacen (1) se denomina **conjunto solución** del sistema (1). A menudo conviene relacionar este conjunto directamente con la matriz A y la ecuación $A\mathbf{x} = \mathbf{0}$. Al conjunto de las $\mathbf{x}$ que satisfacen $A\mathbf{x} = 0$ se le llamará **espacio nulo** de la matriz A.

DEFINICIÓN El **espacio nulo** de una matriz A de $m \times n$, que se escribe Nul A, es el conjunto de todas las soluciones de la ecuación homogénea $A\mathbf{x} = \mathbf{0}$. En notación de conjuntos,

$$\text{Nul } A = \{\mathbf{x} : \mathbf{x} \text{ está en } \mathbf{R}^n \text{ y } A\mathbf{x} = \mathbf{0}\}$$

Una descripción más dinámica de Nul A es el conjunto de todas las $\mathbf{x}$ en $\mathbb{R}^n$ que se mapean en el vector cero de $\mathbb{R}^m$ mediante la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$. Vea la figura 1.

EJEMPLO 1 Sea A como en (2), y sea $\mathbf{u} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$. Determine si $\mathbf{u}$ pertenece al espacio nulo de A.

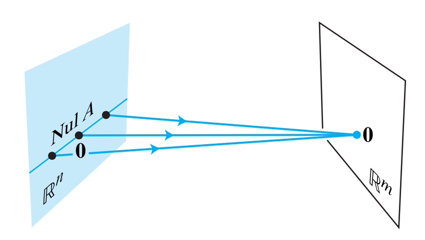


FIGURA 1

Solución Para probar si **u** satisface $A\mathbf{u} = \mathbf{0}$, simplemente calcule

$$A\mathbf{u} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 - 9 + 4 \\ -25 + 27 - 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Entonces **u** está en Nul A. ■

El término espacio en *espacio nulo* resulta adecuado porque el espacio nulo de una matriz es un espacio vectorial, como se verá en el teorema siguiente.

TEOREMA 2

El espacio nulo de una matriz A de $m \times n$ es un subespacio de $\mathbb{R}^n$. De manera equivalente, el conjunto de todas las soluciones de un sistema $A\mathbf{x} = \mathbf{0}$ de m ecuaciones lineales homogéneas con n incógnitas es un subespacio de $\mathbb{R}^n$.

DEMOSTRACIÓN Resulta evidente que Nul A es un subconjunto de $\mathbb{R}^n$ porque A tiene n columnas. Se debe mostrar que Nul A satisface las tres propiedades de un subespacio. Desde luego, **0** está en Nul A. Enseguida, sean **u** y **v** dos vectores cualesquiera de Nul A. Entonces

$$A\mathbf{u} = \mathbf{0} \quad \text{y} \quad A\mathbf{v} = \mathbf{0}$$

Para mostrar que $\mathbf{u} + \mathbf{v}$ está en Nul A, debe probarse que $A(\mathbf{u} + \mathbf{v}) = \mathbf{0}$. Mediante el uso de una propiedad de la multiplicación de matrices, se encuentra que

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Entonces $\mathbf{u} + \mathbf{v}$ está en Nul A, y Nul A es cerrado bajo la suma de vectores. Por último, si c es cualquier escalar, entonces

$$A(c\mathbf{u}) = c(A\mathbf{u}) = c(\mathbf{0}) = \mathbf{0}$$

lo cual demuestra que $c\mathbf{u}$ está en Nul A. Entonces Nul A es un subespacio de $\mathbb{R}^n$. ■

EJEMPLO 2 Sea H el conjunto de todos los vectores en $\mathbb{R}^4$ cuyas coordenadas a, b, c, d satisfacen las ecuaciones $a - 2b + 5c = d$ y $c - a = b$. Demuestre que H es un subespacio de $\mathbb{R}^4$.

Solución Al reacomodar las ecuaciones que describen los elementos de H, se observa que H es el conjunto de todas las soluciones del siguiente sistema homogéneo de ecuaciones lineales:

$$\begin{aligned} a - 2b + 5c - d &= 0 \\ -a - b + c \quad\quad &= 0 \end{aligned}$$

Por el teorema 2, H es un subespacio de $\mathbb{R}^4$.

Es importante que las ecuaciones lineales que definen al conjunto H sean homogéneas. En caso contrario, definitivamente, el conjunto de soluciones *no* será un subespacio (puesto que el vector cero no es solución de un sistema no homogéneo). También, en algunos casos, el conjunto de las soluciones podría estar vacío.

Una descripción explícita de Nul A

No hay ninguna relación evidente entre los vectores de Nul A y las entradas de A. Se dice que Nul A está definido *implícitamente*, porque se define mediante una condición que debe verificarse. No hay una descripción ni una lista explícita de los elementos contenidos en Nul A. Sin embargo, *resolver* la ecuación $A\mathbf{x} = \mathbf{0}$ equivale a producir una descripción *explícita* de Nul A. En el siguiente ejemplo, se repasará el procedimiento de la sección 1.5.

EJEMPLO 3 Encuentre un conjunto generador para el espacio nulo de la matriz

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

Solución El primer paso es encontrar la solución general de $A\mathbf{x} = \mathbf{0}$ en términos de variables libres. Reduzca por filas la matriz aumentada $[A \quad \mathbf{0}]$ a la forma escalonada *reducida* para escribir las variables básicas en términos de las variables libres:

$$\begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 - 2x_2 \quad - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

La solución general es $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, con x_2, x_4 y x_5 libres. Enseguida, descomponga el vector que proporciona la solución general como una combinación lineal de vectores, donde *los pesos son las variables libres.* Esto es,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\substack{\uparrow \\ \mathbf{w}}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}}$$

$$= x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w} \tag{3}$$

Toda combinación lineal de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ es un elemento de Nul A. Entonces $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es un conjunto generador para Nul A.

Es necesario mencionar dos aspectos de la solución del ejemplo 3 que son aplicables a todos los problemas de este tipo. Estos hechos se usarán posteriormente.

1. El conjunto generador producido por el método del ejemplo 3 es, en forma automática, linealmente independiente puesto que las variables libres son los pesos de los vectores generadores. Por ejemplo, observe las entradas segunda, cuarta y quinta del vector solución en (3) y advierta que $x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$ puede ser $\mathbf{0}$ sólo si los pesos x_2, x_4 y x_5 son todos cero.

2. El número de vectores presentes en el conjunto generador para Nul A es igual al número de variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$.

El espacio columna de una matriz

Otro subespacio importante asociado a una matriz es su espacio columna. A diferencia del espacio nulo, el espacio columna se define explícitamente por medio de combinaciones lineales.

DEFINICIÓN El **espacio columna** de una matriz A de $m \times n$, se escribe Col A, es el conjunto de todas las combinaciones lineales de las columnas de A. Si $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$, entonces

$$\text{Col } A = \text{Gen}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$$

Como Gen$\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ es un subespacio, por el teorema 1, el teorema siguiente proviene de la definición de Col A y de que las columnas de A están en $\mathbb{R}^m$.

TEOREMA 3 El espacio columna de una matriz A de $m \times n$ es un subespacio de $\mathbb{R}^m$.

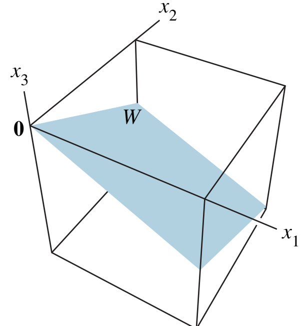
Observe que un vector típico en Col A puede escribirse como $A\mathbf{x}$ para alguna $\mathbf{x}$, porque la notación $A\mathbf{x}$ representa una combinación lineal de columnas de A. Esto es,

$$\boxed{\text{Col } A = \{\mathbf{b} : \mathbf{b} = A\mathbf{x} \text{ para alguna } \mathbf{x} \text{ en } \mathbb{R}^n\}}$$

La notación $A\mathbf{x}$ para vectores en Col A también muestra que Col A es el *rango* de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$. Al final de la sección se retomará este punto de vista.

EJEMPLO 4 Encuentre una matriz A tal que $W = \text{Col } A$.

$$W = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \text{ en } \mathbb{R} \right\}$$



Solución Primero, escriba W como un conjunto de combinaciones lineales.

$$W = \left\{ a \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a, b \text{ en } \mathbb{R} \right\} = \text{Gen}\left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

En segundo lugar, use los vectores del conjunto generador como columnas de A. Sea $A = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ -7 & 0 \end{bmatrix}$. Entonces $W = \text{Col}\, A$, tal como se deseaba. ■

Recuerde que, por el teorema 4 de la sección 1.4, las columnas de A generan $\mathbb{R}^m$ si, y sólo si, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$. Este hecho puede replantearse de la siguiente forma:

El espacio columna de una matriz A de $m \times n$ es todo $\mathbb{R}^m$ si, y sólo si, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^m$.

El contraste entre Nul A y Col A

Es natural preguntarse cómo están relacionados el espacio nulo y el espacio columna de una matriz. De hecho, estos dos espacios son muy diferentes, como se mostrará en los ejemplos 5, 6 y 7. Sin embargo, existe una conexión sorprendente entre ambos que surgirá en la sección 4.6, cuando se tenga un poco más de teoría.

EJEMPLO 5 Sea

$$A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$$

a. Si el espacio columna de A es un subespacio de $\mathbb{R}^k$, ¿cuál es el valor de k?
b. Si el espacio nulo de A es un subespacio de $\mathbb{R}^k$, ¿cuál es el valor de k?

Solución

a. Cada columna de A tiene tres entradas, así que Col A es un subespacio de $\mathbb{R}^k$, donde $k = 3$.
b. Un vector $\mathbf{x}$ tal que $A\mathbf{x}$ esté definido debe tener cuatro entradas, así que Nul A es un subespacio de $\mathbb{R}^k$, donde $k = 4$. ■

Cuando una matriz no es cuadrada, como en el ejemplo 5, los vectores de Nul A y Col A tienen lugar en "universos" completamente diferentes. Por ejemplo, ninguna combinación lineal de vectores en $\mathbb{R}^3$ puede producir un vector en $\mathbb{R}^4$. Cuando A es cuadrada, Nul A y Col A tienen el vector cero en común, y en casos especiales es posible que algunos vectores diferentes de cero pertenezcan tanto a Nul A como a Col A.

EJEMPLO 6 Con A igual que en el ejemplo 5, encuentre un vector distinto de cero en Col A y un vector distinto de cero en Nul A.

Solución Es fácil encontrar un vector distinto de cero en Col A. Cualquier columna de A sirve, por ejemplo, $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$. Para encontrar un vector distinto de cero en Nul A, es necesario trabajar un poco. Reduzca por filas la matriz aumentada $[A \quad \mathbf{0}]$ para obtener

$$[A \quad \mathbf{0}] \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Así que, si $\mathbf{x}$ satisface $A\mathbf{x} = \mathbf{0}$, entonces $x_1 = -9x_3$, $x_2 = 5x_3$, $x_4 = 0$, y x_3 es libre. Al asignar un valor distinto de cero a x_3 —por ejemplo, $x_3 = 1$— se obtiene un vector en Nul A, a saber, $\mathbf{x} = (-9, 5, 1, 0)$.

EJEMPLO 7 Con la A del ejemplo 5, sean $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

a. Determine si $\mathbf{u}$ está en Nul A. ¿Podría $\mathbf{u}$ estar en Col A?

b. Determine si $\mathbf{v}$ está en Col A. ¿Podría $\mathbf{v}$ estar en Nul A?

Solución

a. En este momento no se necesita una descripción explícita de Nul A. Simplemente calcule el producto $A\mathbf{u}$.

$$A\mathbf{u} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Desde luego, $\mathbf{u}$ *no* es solución de $A\mathbf{x} = \mathbf{0}$, así que $\mathbf{u}$ no está en Nul A. Además, con cuatro entradas, $\mathbf{u}$ no podría estar en Col A, puesto que Col A es un subespacio de $\mathbb{R}^3$.

b. Reduzca $[A \quad \mathbf{v}]$ a una forma escalonada.

$$[A \quad \mathbf{v}] = \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ -2 & -5 & 7 & 3 & -1 \\ 3 & 7 & -8 & 6 & 3 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ 0 & 1 & -5 & -4 & -2 \\ 0 & 0 & 0 & 17 & 1 \end{bmatrix}$$

En este punto, resulta claro que la ecuación $A\mathbf{x} = \mathbf{v}$ es consistente, así que $\mathbf{v}$ está en Col A. Con sólo tres entradas, $\mathbf{v}$ no podría estar en Nul A, puesto que Nul A es un subespacio de $\mathbb{R}^4$.

La tabla siguiente resume lo aprendido hasta ahora acerca de Nul A y Col A. El punto 8 es una reformulación de los teoremas 11 y 12(a) de la sección 1.9.

Contraste entre Nul A y Col A para una matriz A $m \times n$

Nul A	Col A
1. Nul A es un subespacio de $\mathbb{R}^n$.	**1.** Col A es un subespacio de $\mathbb{R}^m$.
2. Nul A está definido implícitamente; esto es, sólo se tiene una condición ($A\mathbf{x} = \mathbf{0}$) que los vectores de Nul A deben satisfacer.	**2.** Col A está definido explícitamente; esto es, se especifica cómo construir los vectores de Col A.
3. Se requiere tiempo para encontrar vectores en Nul A. Son necesarias operaciones por fila con $[A \quad \mathbf{0}]$.	**3.** Es fácil encontrar los vectores de Col A. Las columnas de A se despliegan, y se forman otras columnas a partir de ellas.
4. No hay una relación evidente entre Nul A y las entradas de A.	**4.** Hay una relación evidente entre Col A y las entradas de A, puesto que cada columna de A está en Col A.
5. Un vector típico $\mathbf{v}$ en Nul A tiene la propiedad de que $A\mathbf{v} = \mathbf{0}$.	**5.** Un vector típico $\mathbf{v}$ en Col A tiene la propiedad de que la ecuación $A\mathbf{x} = \mathbf{v}$ es consistente.
6. Dado un vector específico $\mathbf{v}$, es fácil saber si $\mathbf{v}$ está en Nul A. Sólo calcule $A\mathbf{v}$.	**6.** Dado un vector específico $\mathbf{v}$, puede tomar algún tiempo decidir si $\mathbf{v}$ está en Col A. Se necesitan operaciones por fila sobre $[A \quad \mathbf{v}]$.
7. Nul $A = \{\mathbf{0}\}$ si, y sólo si, la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial.	**7.** Col $A = \mathbb{R}^m$ si, y sólo si, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene solución para cada $\mathbf{b}$ en $\mathbb{R}^m$.
8. Nul $A = \{\mathbf{0}\}$ si, y sólo si, la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno.	**8.** Col $A = \mathbb{R}^m$ si, y sólo si, la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^n$ *sobre* $\mathbb{R}^m$.

Núcleo y rango de una transformación lineal

Los subespacios de espacios vectoriales distintos de $\mathbb{R}^n$ a menudo se describen en términos de transformaciones lineales, en vez de una matriz. Para precisar esto, se generalizará la definición dada en la sección 1.8.

DEFINICIÓN

Una **transformación lineal** T de un espacio vectorial V a un espacio vectorial W es una regla que asigna a cada vector $\mathbf{x}$ en V un único vector $T(\mathbf{x})$ en W, de modo que

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para todos $\mathbf{u}$, $\mathbf{v}$ en V, y

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ para todo $\mathbf{u}$ en V y todos los escalares c.

El **núcleo** (o **espacio nulo**) de una T como la anterior es el conjunto de todos los $\mathbf{u}$ en V tales que $T(\mathbf{u}) = \mathbf{0}$ (el vector cero en W). El **rango** de T es el conjunto de todos los vectores en W de la forma $T(\mathbf{x})$ para alguna $\mathbf{x}$ en V. Si resulta que T proviene de una transformación matricial —por ejemplo, $T(\mathbf{x}) = A\mathbf{x}$ para alguna matriz A—, entonces el núcleo y el rango de T son simplemente el espacio nulo y el espacio columna de A, tal como se definieron antes.

No es difícil demostrar que el núcleo de T es un subespacio de V. La demostración es esencialmente la misma que la del teorema 2. También, el rango de T es un subespacio de W. Vea la figura 2 y el ejercicio 30.

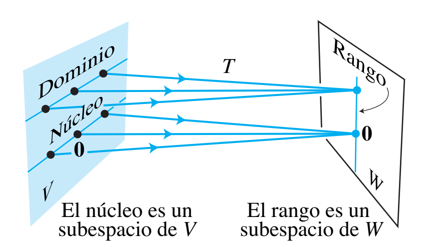


FIGURA 2 Subespacios asociados con una transformación lineal.

En las aplicaciones, un subespacio suele surgir como el núcleo o el rango de una transformación lineal adecuada. Por ejemplo, el conjunto de todas las soluciones de una ecuación diferencial lineal homogénea resulta ser el núcleo de una transformación lineal. De manera característica, una transformación lineal de este tipo se describe en términos de una o más derivadas de una función. Explicar esto con todo detalle significaría alejarse demasiado del tema principal en este momento, así que sólo se presentarán dos ejemplos. El primer ejemplo explica por qué la operación de diferenciación es una transformación lineal.

EJEMPLO 8 (*Se requiere cálculo.*) Sea V el espacio vectorial de todas las funciones f con valores reales definidas sobre un intervalo $[a, b]$, con la propiedad de que son diferenciables y sus derivadas son continuas en $[a, b]$. Sea W el espacio vectorial de todas las funciones continuas en $[a, b]$, y sea $D : V \to W$ la transformación que convierte a f en V en su derivada f'. En cálculo, dos de las sencillas reglas de diferenciación son

$$D(f + g) = D(f) + D(g) \quad \text{y} \quad D(cf) = cD(f)$$

Esto es, D es una transformación lineal. Se puede mostrar que el núcleo de D es el conjunto de funciones constantes sobre [a, b], y que el rango de D es el conjunto W de todas las funciones continuas sobre $[a, b]$.

EJEMPLO 9 (*Se requiere cálculo.*) La ecuación diferencial

$$y'' + \omega^2 y = 0 \tag{4}$$

donde ω es una constante, se usa para describir diversos sistemas físicos, como la vibración de un resorte unido a un peso, el movimiento de un péndulo, y el voltaje en un circuito eléctrico con inductancia y capacitancia. El conjunto de soluciones dado en (4) es precisamente el núcleo de la transformación lineal que mapea una función $y = f(t)$ en una función $f''(t) + \omega^2 f(t)$. Encontrar una descripción explícita de este espacio vectorial es un problema de ecuaciones diferenciales. La solución resulta ser el espacio descrito en el ejercicio 19 de la sección 4.1.

PROBLEMAS DE PRÁCTICA

1. Sea $W = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a - 3b - c = 0 \right\}$. Muestre que W es un subespacio de $\mathbb{R}^3$ en dos formas diferentes. (Use dos teoremas.)

2. Sean $A = \begin{bmatrix} 7 & -3 & 5 \\ -4 & 1 & -5 \\ -5 & 2 & -4 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix}$. Suponga que sabe que las ecuaciones $A\mathbf{x} = \mathbf{v}$ y $A\mathbf{x} = \mathbf{w}$ son consistentes. ¿Qué puede decirse acerca de la ecuación $A\mathbf{x} = \mathbf{v} + \mathbf{w}$?

4.2 EJERCICIOS

1. Determine si $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$ está en Nul A, donde

$$A = \begin{bmatrix} 3 & -5 & -3 \\ 6 & -2 & 0 \\ -8 & 4 & 1 \end{bmatrix}.$$

2. Determine si $\mathbf{w} = \begin{bmatrix} 5 \\ -3 \\ 2 \end{bmatrix}$ está en Nul A, donde

$$A = \begin{bmatrix} 5 & 21 & 19 \\ 13 & 23 & 2 \\ 8 & 14 & 1 \end{bmatrix}.$$

En los ejercicios 3 a 6, encuentre una descripción explícita de Nul A, para ello enliste los vectores que generan el espacio nulo.

3. $A = \begin{bmatrix} 1 & 3 & 5 & 0 \\ 0 & 1 & 4 & -2 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & -6 & 4 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & -2 & 0 & 4 & 0 \\ 0 & 0 & 1 & -9 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 5 & -4 & -3 & 1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 7 a 14, use un teorema adecuado para mostrar que el conjunto dado, W, es un espacio vectorial, o encuentre un ejemplo específico de lo contrario.

7. $\left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a + b + c = 2 \right\}$

8. $\left\{ \begin{bmatrix} r \\ s \\ t \end{bmatrix} : 5r - 1 = s + 2t \right\}$

9. $\left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} : \begin{matrix} a - 2b = 4c \\ 2a = c + 3d \end{matrix} \right\}$

10. $\left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} : \begin{matrix} a + 3b = c \\ b + c + a = d \end{matrix} \right\}$

11. $\left\{ \begin{bmatrix} b - 2d \\ 5 + d \\ b + 3d \\ d \end{bmatrix} : b, d \text{ reales} \right\}$

12. $\left\{ \begin{bmatrix} b - 5d \\ 2b \\ 2d + 1 \\ d \end{bmatrix} : b, d \text{ reales} \right\}$

13. $\left\{ \begin{bmatrix} c - 6d \\ d \\ c \end{bmatrix} : c, d \text{ reales} \right\}$

14. $\left\{ \begin{bmatrix} -a + 2b \\ a - 2b \\ 3a - 6b \end{bmatrix} : a, b \text{ reales} \right\}$

En los ejercicios 15 y 16, encuentre una A tal que el conjunto dado sea Col A.

15. $\left\{ \begin{bmatrix} 2s + 3t \\ r + s - 2t \\ 4r + s \\ 3r - s - t \end{bmatrix} : r, s, t \text{ reales} \right\}$

16. $\left\{ \begin{bmatrix} b - c \\ 2b + c + d \\ 5c - 4d \\ d \end{bmatrix} : b, c, d \text{ reales} \right\}$

Para las matrices de los ejercicios 17 a 20, (a) encuentre una k tal que Nul A sea un subespacio de $\mathbb{R}^k$, y (b) encuentre una k tal que Col A sea un subespacio de $\mathbb{R}^k$.

17. $A = \begin{bmatrix} 2 & -6 \\ -1 & 3 \\ -4 & 12 \\ 3 & -9 \end{bmatrix}$

18. $A = \begin{bmatrix} 7 & -2 & 0 \\ -2 & 0 & -5 \\ 0 & -5 & 7 \\ -5 & 7 & -2 \end{bmatrix}$

19. $A = \begin{bmatrix} 4 & 5 & -2 & 6 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$

20. $A = \begin{bmatrix} 1 & -3 & 9 & 0 & -5 \end{bmatrix}$

21. Con A como en el ejercicio 17, encuentre un vector distinto de cero en Nul A y un vector distinto de cero en Col A.

22. Con A como en el ejercicio 3, encuentre un vector distinto de cero en Nul A y un vector distinto de cero en Col A.

23. Sea $A = \begin{bmatrix} -6 & 12 \\ -3 & 6 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Determine si $\mathbf{w}$ está en Col A. ¿Está $\mathbf{w}$ en Nul A?

24. Sea $A = \begin{bmatrix} -8 & -2 & -9 \\ 6 & 4 & 8 \\ 4 & 0 & 4 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$. Determine si $\mathbf{w}$ está en Col A. ¿Está $\mathbf{w}$ en Nul A?

En los ejercicios 25 y 26, A denota una matriz de $m \times n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

25. a. El espacio nulo de A es el conjunto solución de la ecuación $A\mathbf{x} = \mathbf{0}$.

b. El espacio nulo de una matriz de $m \times n$ está en $\mathbb{R}^m$.

c. El espacio columna de A es el rango de la función $\mathbf{x} \mapsto A\mathbf{x}$.

d. Si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces Col A es $\mathbb{R}^m$.

e. El núcleo de una transformación lineal es un espacio vectorial.

f. Col A es el conjunto de todos los vectores que pueden escribirse como $A\mathbf{x}$ para alguna $\mathbf{x}$.

26. a. Un espacio nulo es un espacio vectorial.

b. El espacio columna de una matriz de $m \times n$ está en $\mathbb{R}^m$.

c. Col A es el conjunto de todas las soluciones de $A\mathbf{x} = \mathbf{b}$.

d. Nul A es el núcleo de la función $\mathbf{x} \mapsto A\mathbf{x}$.

e. El rango de una transformación lineal es un espacio vectorial.

f. El conjunto de todas las soluciones de una ecuación diferencial lineal homogénea es el núcleo de una transformación lineal.

27. Puede mostrarse que una solución del sistema siguiente es $x_1 = 3$, $x_2 = 2$, y $x_3 = -1$. Use este hecho y la teoría de esta sección para explicar por qué otra solución es $x_1 = 30$, $x_2 = 20$, y $x_3 = -10$. (Observe cómo están relacionadas las soluciones, pero no realice otros cálculos.)

$$\begin{aligned} x_1 - 3x_2 - 3x_3 &= 0 \\ -2x_1 + 4x_2 + 2x_3 &= 0 \\ -x_1 + 5x_2 + 7x_3 &= 0 \end{aligned}$$

28. Considere los dos sistemas de ecuaciones siguientes:

$$\begin{aligned} 5x_1 + x_2 - 3x_3 &= 0 \\ -9x_1 + 2x_2 + 5x_3 &= 1 \\ 4x_1 + x_2 - 6x_3 &= 9 \end{aligned} \qquad \begin{aligned} 5x_1 + x_2 - 3x_3 &= 0 \\ -9x_1 + 2x_2 + 5x_3 &= 5 \\ 4x_1 + x_2 - 6x_3 &= 45 \end{aligned}$$

Se puede comprobar que el primer sistema tiene solución. Use este hecho y la teoría de esta sección para explicar por qué el segundo sistema también debe tener solución. (No haga operaciones por fila.)

29. Demuestre el teorema 3 de la siguiente manera: dada un matriz A de $m \times n$, un elemento de Col A tiene la forma $A\mathbf{x}$ para alguna $\mathbf{x}$ en $\mathbb{R}^n$. Sean $A\mathbf{x}$ y $A\mathbf{w}$ cualesquiera dos vectores incluidos en Col A.

a. Explique por qué el vector cero está en Col A.

b. Muestre que el vector $A\mathbf{x} + A\mathbf{w}$ está en Col A.

c. Dado un escalar c, demuestre que $c(A\mathbf{x})$ está en Col A.

30. Sea $T: V \to W$ una transformación lineal de un espacio vectorial V en un espacio vectorial W. Demuestre que el rango de T es un subespacio de W. [*Sugerencia:* Los elementos típicos del rango tienen la forma $T(\mathbf{x})$ y $T(\mathbf{w})$ para algunas $\mathbf{x}$, $\mathbf{w}$ en V.]

31. Defina $T: \mathbb{P}_2 \to \mathbb{R}^2$ por medio de $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$. Por ejemplo, si $\mathbf{p}(t) = 3 + 5t + 7t^2$, entonces $T(\mathbf{p}) = \begin{bmatrix} 3 \\ 15 \end{bmatrix}$.

a. Muestre que T es una transformación lineal. [*Sugerencia:* Para polinomios arbitrarios $\mathbf{p}$ y $\mathbf{q}$ en $\mathbb{P}_2$, calcule $T(\mathbf{p} + \mathbf{q})$ y $T(c\mathbf{p})$.]

b. Encuentre un polinomio $\mathbf{p}$ en $\mathbb{P}_2$ que genere el núcleo de T, y describa el rango de T.

32. Defina una transformación lineal $T: \mathbb{P}_2 \to \mathbb{R}^2$ por medio de $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(0) \end{bmatrix}$. Encuentre los polinomios $\mathbf{p}_1$ y $\mathbf{p}_2$ en $\mathbb{P}_2$ que generen el núcleo de T, y describa el rango de T.

33. Sea $M_{2\times 2}$ el espacio vectorial de todas las matrices de 2×2, y defina $T: M_{2\times 2} \to M_{2\times 2}$ como $T(A) = A + A^T$, donde $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

a. Muestre que T es una transformación lineal.

b. Sea B cualquier elemento de $M_{2\times 2}$ tal que $B^T = B$. Encuentre una A en $M_{2\times 2}$ tal que $T(A) = B$.

c. Muestre que el rango de T es el conjunto B en $M_{2\times 2}$ con la propiedad de que $B^T = B$.

d. Describa el núcleo de T.

34. (*Se requiere cálculo.*) Defina $T : C[0, 1] \to C[0, 1]$ de la siguiente forma: para $\mathbf{f}$ en $C[0, 1]$, sea $T(\mathbf{f})$ la antiderivada $\mathbf{F}$ de $\mathbf{f}$ tal que $\mathbf{F}(0) = 0$. Demuestre que T es una transformación lineal, y describa el núcleo de T. (Vea la notación dada en el ejercicio 20 de la sección 4.1.)

35. Sean V y W dos espacios vectoriales, y sea $T : V \to W$ una transformación lineal. Dado un subespacio U de V, denote con $T(U)$ el conjunto de imágenes de la forma $T(\mathbf{x})$, donde $\mathbf{x}$ está en U. Demuestre que $T(U)$ es un subespacio de W.

36. Dada $T : V \to W$ como en el ejercicio 35, y dado un subespacio Z de W, sea U el conjunto de todas las $\mathbf{x}$ en V tal que $T(\mathbf{x})$ esté en Z. Muestre que U es un subespacio de V.

37. **[M]** Determine si $\mathbf{w}$ está en el espacio columna de A, en el espacio nulo de A, o en ambos, donde

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -3 \end{bmatrix}, \quad A = \begin{bmatrix} 7 & 6 & -4 & 1 \\ -5 & -1 & 0 & -2 \\ 9 & -11 & 7 & -3 \\ 19 & -9 & 7 & 1 \end{bmatrix}$$

38. **[M]** Determine si $\mathbf{w}$ está en el espacio columna de A, en el espacio nulo de A, o en ambos, donde

$$\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \quad A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$$

39. **[M]** Sean $\mathbf{a}_1, \ldots, \mathbf{a}_5$ las columnas de la matriz A, donde

$$A = \begin{bmatrix} 5 & 1 & 2 & 2 & 0 \\ 3 & 3 & 2 & -1 & -12 \\ 8 & 4 & 4 & -5 & 12 \\ 2 & 1 & 1 & 0 & -2 \end{bmatrix}, \quad B = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_4\,]$$

a. Explique por qué $\mathbf{a}_3$ y $\mathbf{a}_5$ están en el espacio columna de B.

b. Encuentre un conjunto de vectores que genere Nul A.

c. Sea $T : \mathbb{R}^5 \to \mathbb{R}^4$ definida mediante $T(\mathbf{x}) = A\mathbf{x}$. Explique por qué T no es inyectiva ni suprayectiva.

40. **[M]** Sean $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$ y $K = \text{Gen}\{\mathbf{v}_3, \mathbf{v}_4\}$, donde

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ 3 \\ 8 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 \\ -12 \\ -28 \end{bmatrix}.$$

Entonces H y K son subespacios de $\mathbb{R}^3$. De hecho, H y K son planos en $\mathbb{R}^3$ que pasan por el origen y se intersecan en una línea que pasa por $\mathbf{0}$. Encuentre un vector $\mathbf{w}$ distinto de cero que genere dicha línea. [*Sugerencia:* $\mathbf{w}$ puede escribirse como $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, y también como $c_3\mathbf{v}_3 + c_4\mathbf{v}_4$. Para construir $\mathbf{w}$, resuelva la ecuación $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_3\mathbf{v}_3 + c_4\mathbf{v}_4$ para las incógnitas c_j.]

SG Dominio de espacios vectoriales, subespacios, Col A y Nul A 4 a 7 (Mastering: Vector Space, Subspace, Col A, and Nul A 4-7)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. *Primer método:* W es un subespacio de $\mathbb{R}^3$ porque, según el teorema 2, es el conjunto de todas las soluciones de un sistema homogéneo de ecuaciones lineales (donde el sistema consta sólo de una ecuación). De manera equivalente, W es el espacio nulo de la matriz de 1×3 $A = [1 \quad -3 \quad -1]$. *Segundo método:* Resuelva la ecuación $a - 3b - c = 0$ para la variable delantera a en términos de las variables libres b y c. Cualquier solución tiene la forma $\begin{bmatrix} 3b + c \\ b \\ c \end{bmatrix}$, donde b y c son arbitrarios, y

$$\begin{bmatrix} 3b + c \\ b \\ c \end{bmatrix} = b \underset{\uparrow \atop \mathbf{v}_1}{\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}} + c \underset{\uparrow \atop \mathbf{v}_2}{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}$$

Este cálculo muestra que $W = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$. Entonces W es un subespacio de $\mathbb{R}^3$ según el teorema 1. También podría resolverse la ecuación $a - 3b - c = 0$ para b o c, y así obtener descripciones alternativas de W como un conjunto de combinaciones lineales de dos vectores.

2. Tanto **v** como **w** están en Col A. Puesto que Col A es un espacio vectorial, $\mathbf{v} + \mathbf{w}$ debe estar en Col A. Esto es, la ecuación $A\mathbf{x} = \mathbf{v} + \mathbf{w}$ es consistente.

4.3 CONJUNTOS LINEALMENTE INDEPENDIENTES; BASES

En esta sección se identificarán y estudiarán los subconjuntos que generan un espacio vectorial V o un subespacio H de la manera más "eficiente" posible. La idea clave es la de independencia lineal, definida como en $\mathbb{R}^n$.

Se dice que un conjunto indexado de vectores $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en V es **linealmente independiente** si la ecuación vectorial

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \tag{1}$$

tiene *solamente* la solución trivial, $c_1 = 0, \ldots, c_p = 0$.[1]

Se dice que el conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es **linealmente dependiente** si (1) tiene una solución no trivial, esto es, si existen pesos $c_1, \ldots, c_p$, *no todos cero*, de modo que se cumpla (1). En tal caso, se afirma que (1) es una **relación de dependencia lineal** entre $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

Igual que en $\mathbb{R}^n$, un conjunto que contiene un único vector **v** es linealmente independiente si, y sólo si, $\mathbf{v} \neq \mathbf{0}$. Además, un conjunto con dos vectores es linealmente dependiente si, y sólo si, uno de los vectores es múltiplo del otro. Cualquier conjunto que contenga el vector cero es linealmente dependiente. El teorema siguiente tiene la misma demostración que el teorema 7 de la sección 1.7.

TEOREMA 4

Un conjunto indexado $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ de dos o más vectores, con $\mathbf{v}_1 \neq \mathbf{0}$, es linealmente dependiente si, y sólo si, algún $\mathbf{v}_j$ (con $j > 1$) es una combinación lineal de los vectores anteriores, $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$.

La principal diferencia entre la dependencia lineal en $\mathbb{R}^n$ y en un espacio vectorial general es que, cuando los vectores no son n-adas, la ecuación homogénea (1) usualmente no se puede escribir como un sistema de n ecuaciones lineales. Esto es, los vectores no pueden colocarse como las columnas de una matriz A para estudiar la ecuación $A\mathbf{x} = 0$. En vez de ello, es necesario recurrir a la definición de dependencia lineal y al teorema 4.

EJEMPLO 1 Sean $\mathbf{p}_1(t) = 1$, $\mathbf{p}_2(t) = t$, y $\mathbf{p}_3(t) = 4 - t$. Entonces $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ es linealmente dependiente en $\mathbb{P}$ porque $\mathbf{p}_3 = 4\mathbf{p}_1 - \mathbf{p}_2$.

EJEMPLO 2 El conjunto $\{\text{sen } t, \cos t\}$ es linealmente independiente en $C[0, 1]$, el espacio de todas las funciones continuas sobre $0 \leq t \leq 1$, porque sen t y cos t no son múltiplos uno del otro *como vectores en* $C[0, 1]$. Esto es, no existe un escalar c tal que

[1] Resulta conveniente usar en (1) $c_1, \ldots, c_p$ para denotar los escalares, en lugar de $x_1, \ldots, x_p$ como se hizo en el capítulo 1.

$\cos t = c \cdot \text{sen}\, t$ para toda t en [0, 1]. (Vea las gráficas de sen t y cos t.) Sin embargo, {sen t cos t, sen $2t$} es linealmente dependiente a causa de la identidad: sen $2t = 2$ sen t cos t, para toda t.

DEFINICIÓN Sea H un subespacio de un espacio vectorial V. Un conjunto indexado de vectores $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ en V es una **base** para H si

(i) $\mathcal{B}$ es un conjunto linealmente independiente, y

(ii) el subespacio generado por $\mathcal{B}$ coincide con H, esto es,

$$H = \text{Gen}\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$$

La definición de base se aplica al caso en que $H = V$, porque cualquier espacio vectorial es un subespacio de sí mismo. Así, una base de V es un conjunto linealmente independiente que genera V. Observe: cuando $H \neq V$, la condición (ii) incluye el requisito de que cada uno de los vectores $\mathbf{b}_1, \ldots, \mathbf{b}_p$ debe pertenecer a H, porque Gen$\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ contiene a $\mathbf{b}_1, \ldots, \mathbf{b}_p$, tal como se vio en la sección 4.1.

EJEMPLO 3 Sea A una matriz invertible de $n \times n$, por ejemplo, $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$. Entonces las columnas de A forman una base para $\mathbb{R}^n$ porque son linealmente independientes y generan $\mathbb{R}^n$, de acuerdo con el teorema de la matriz invertible.

EJEMPLO 4 Sean $\mathbf{e}_1, \ldots, \mathbf{e}_n$ las columnas de la matriz identidad de $n \times n$, I_n. Esto es

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

El conjunto $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ se denomina **base estándar** de $\mathbb{R}^n$ (figura 1).

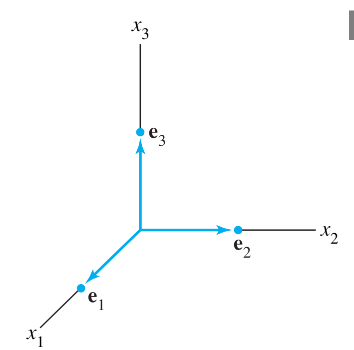


FIGURA 1
La base estándar para $\mathbb{R}^3$.

EJEMPLO 5 Sean $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}$, y $\mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$. Determine si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base para $\mathbb{R}^3$.

Solución Dado que existen exactamente tres vectores en $\mathbb{R}^3$, puede usarse alguno de los diversos métodos para determinar si la matriz $A = [\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \mathbf{v}_3]$ es invertible. Por ejemplo, con dos reemplazos de fila se revela que A tiene tres posiciones pivote. Entonces A es invertible. Al igual que en el ejemplo 3, las columnas de A forman una base para $\mathbb{R}^3$.

EJEMPLO 6 Sea $S = \{1, t, t^2, \ldots, t^n\}$. Verifique que S es una base para $\mathbb{P}_n$. Esta base es llamada **base estándar** para $\mathbb{P}_n$.

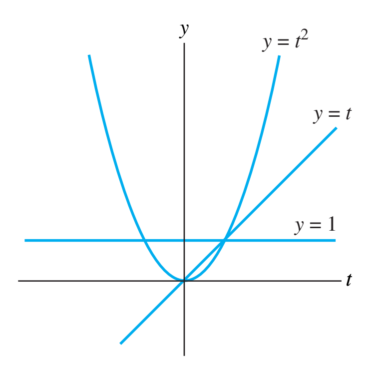


FIGURA 2
La base estándar para $\mathbb{P}_2$.

Solución Desde luego, S genera $\mathbb{P}_n$. Para mostrar que S es linealmente independiente, suponga que $c_0, \ldots, c_n$ satisfacen

$$c_0 \cdot 1 + c_1 t + c_2 t^2 + \cdots + c_n t^n = \mathbf{0}(t) \tag{2}$$

Esta igualdad significa que el polinomio de la izquierda tiene los mismos valores que el polinomio cero de la derecha. Un teorema fundamental del álgebra establece que el único polinomio en $\mathbb{P}_n$ con más de n ceros es el polinomio cero. Esto es, (2) se aplica para toda t sólo si $c_0 = \cdots = c_n = 0$. Esto demuestra que S es linealmente independiente y, por lo tanto, es una base para $\mathbb{P}_n$. Vea la figura 2.

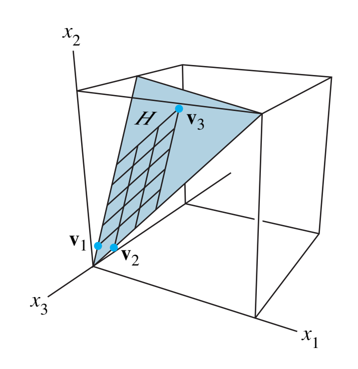
Los problemas que involucran independencia lineal y generación en $\mathbb{P}_n$ se manejan mejor con una técnica que se estudiará en la sección 4.4.

El teorema del conjunto generador

Como se verá, una base es un conjunto generador "eficiente" que no contiene vectores innecesarios. De hecho, se puede construir una base a partir de un conjunto generador descartando algunos vectores innecesarios.

EJEMPLO 7 Sean $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix}$, y $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

Observe que $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, y muestre que $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$. Luego encuentre una base para el subespacio H.



Solución Todo vector en $\text{Gen}(\mathbf{v}_1, \mathbf{v}_2\}$ pertenece a H porque

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + 0\mathbf{v}_3$$

Ahora sea $\mathbf{x}$ cualquier vector en H —por ejemplo, $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Como $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, se puede sustituir

$$\begin{aligned} \mathbf{x} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3(5\mathbf{v}_1 + 3\mathbf{v}_2) \\ &= (c_1 + 5c_3)\mathbf{v}_1 + (c_2 + 3c_3)\mathbf{v}_2 \end{aligned}$$

Entonces $\mathbf{x}$ está en $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2)$, y todo vector en H pertenece ya a $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$. Se concluye que H y $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$ son en realidad el mismo conjunto de vectores. Se deduce que $\{\mathbf{v}_1, \mathbf{v}_2)$ es una base de H pues resulta obvio que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente independiente.

El teorema siguiente generaliza el ejemplo 7.

TEOREMA 5 **Teorema del conjunto generador**

Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ un conjunto en V, y sea $H = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

a. Si uno de los vectores de S —por ejemplo, $\mathbf{v}_k$— es una combinación lineal de los vectores restantes de S, entonces el conjunto que se forma a partir de S al retirarle $\mathbf{v}_k$ todavía genera H.
b. Si $H \neq \{\mathbf{0}\}$, algún subconjunto de S es una base para H.

DEMOSTRACIÓN

a. Al reordenar la lista de vectores de S, si fuera necesario, se puede suponer que $\mathbf{v}_p$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}$ —por ejemplo,

$$\mathbf{v}_p = a_1\mathbf{v}_1 + \cdots + a_{p-1}\mathbf{v}_{p-1} \tag{3}$$

Dada cualquier $\mathbf{x}$ en H, se puede escribir

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_{p-1}\mathbf{v}_{p-1} + c_p\mathbf{v}_p \tag{4}$$

para escalares adecuados $c_1, \ldots, c_p$. Al sustituir la expresión para $\mathbf{v}_p$ de (3) a (4), es fácil advertir que $\mathbf{x}$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}$. Entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}\}$ genera H, porque $\mathbf{x}$ era un elemento arbitrario de H.

b. Si el conjunto original S es linealmente independiente, entonces ya es una base para H. En caso contrario, uno de los vectores de S depende de los otros y puede eliminarse, en vista de (a). Mientras haya dos o más vectores en el conjunto generador, puede repetirse este proceso hasta que el conjunto generador sea linealmente independiente y, por lo tanto, forme una base para H. Si el conjunto generador se reduce finalmente a un vector, ese vector será distinto de cero (y, de esta manera, linealmente independiente) porque $H \neq \{\mathbf{0}\}$. ■

Bases para Nul A y Col A

Ya se sabe cómo encontrar vectores que generen el espacio nulo de una matriz A. En la explicación de la sección 4.2 se afirmó que el método siempre produce un conjunto linealmente independiente. Entonces el método produce una *base* para Nul A.

Los siguientes dos ejemplos describen un algoritmo sencillo con el cual es posible encontrar una base para el espacio columna.

EJEMPLO 8 Encuentre una base para Col B, donde

$$B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Solución Cada columna de B que no es pivote es una combinación lineal de las columnas pivote. De hecho, $\mathbf{b}_2 = 4\mathbf{b}_1$ y $\mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$. Por el teorema del conjunto generador, se pueden desechar $\mathbf{b}_2$ y $\mathbf{b}_4$, y el conjunto $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ todavía generará Col B. Sea

$$S = \{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Como $\mathbf{b}_1 \neq 0$ y ningún vector de S es una combinación lineal de los vectores que le preceden, S es linealmente independiente (teorema 4). Entonces S es una base para Col B. ■

¿Qué sucede con una matriz A que *no* está en forma escalonada reducida? Recuerde que cualquier relación de dependencia lineal entre las columnas de A puede expresarse en la forma $A\mathbf{x} = \mathbf{0}$, donde $\mathbf{x}$ es una columna de pesos. (Si algunas columnas no intervienen en una relación de dependencia dada, sus pesos son cero.) Cuando A se reduce por filas a una matriz B, las columnas de B resultan a menudo totalmente diferentes de las columnas de A. Sin embargo, las ecuaciones $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ tienen exactamente el mismo conjunto de soluciones. Esto es, las columnas de A tienen *exactamente las mismas relaciones de dependencia lineal* que las columnas de B.

> Las operaciones elementales de fila aplicadas a una matriz no afectan las relaciones de dependencia lineal entre las columnas de la matriz.

EJEMPLO 9 Puede mostrarse que la matriz

$$A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_5\,] = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$$

es equivalente por filas a la matriz B del ejemplo 8. Encuentre una base para Col A.

Solución En el ejemplo 8 se vio que

$$\mathbf{b}_2 = 4\mathbf{b}_1 \quad \text{y} \quad \mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$$

entonces puede esperarse que

$$\mathbf{a}_2 = 4\mathbf{a}_1 \quad \text{y} \quad \mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3$$

¡Compruebe que esto realmente sucede! Entonces es posible eliminar $\mathbf{a}_2$ y $\mathbf{a}_4$ al elegir un conjunto generador mínimo para Col A. De hecho, $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ debe ser linealmente independiente porque cualquier relación de dependencia lineal entre $\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5$ implicaría una relación de dependencia lineal entre $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5$. Pero se sabe que $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ es un conjunto linealmente independiente. Entonces $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ es una base para Col A. Las columnas utilizadas para esta base son las columnas pivote de A.

En los ejemplos 8 y 9 se ilustra el siguiente hecho útil.

TEOREMA 6 Las columnas pivote de una matriz A forman una base para Col A.

DEMOSTRACIÓN La demostración general usa los argumentos que acaban de analizarse. Sea B la forma escalonada reducida de A. El conjunto de columnas pivote de B es linealmente independiente, pues ningún vector del conjunto es una combinación lineal de los vectores que lo preceden. Como A es equivalente por filas a B, las columnas pivote de A también son linealmente independientes, porque cualquier relación de dependencia entre las columnas de A corresponde a una relación de dependencia lineal entre las columnas de B. Por esta misma razón, cualquier columna de A que no sea pivote es una combinación lineal de las columnas pivote de A. Entonces las columnas de A que no son pivote

pueden descartarse del conjunto generador para Col A, de acuerdo con el teorema del conjunto generador. Esto deja a las columnas pivote de A como base para Col A. ■

Advertencia: Tenga el cuidado de usar *columnas pivote de la propia A* para la base de Col A. Las columnas de una forma escalonada B de A, a menudo no están en el espacio columna de A. Por ejemplo, todas las columnas de B del ejemplo 8 tienen ceros en su última entrada, así que no pueden generar el espacio columna de la A del ejemplo 9.

Dos perspectivas de una base

Cuando se usa el teorema del conjunto generador, la eliminación de vectores de un conjunto generador debe terminar cuando el conjunto resulta linealmente independiente. Si se elimina un vector adicional, no será una combinación lineal de los vectores restantes, y así el conjunto resultante ya no generará V. Entonces una base es un conjunto generador lo más pequeño posible.

Una base también es un conjunto linealmente independiente lo más grande posible. Si S es una base de V, y si S se amplía con un vector —por ejemplo, $\mathbf{w}$— de V, entonces el nuevo conjunto ya no puede ser linealmente independiente, porque S genera V, y entonces $\mathbf{w}$ es una combinación lineal de los elementos de S.

EJEMPLO 10 Los siguientes tres conjuntos en $\mathbb{R}^3$ muestran cómo un conjunto linealmente independiente puede ampliarse para obtener una base, y cómo una ampliación adicional destruye la independencia lineal del conjunto. También, un conjunto generador puede reducirse para obtener una base, pero una reducción adicional ocasiona que el conjunto ya no sea generador.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix}\right\} \quad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix}\right\} \quad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix},\begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

Linealmente independiente, pero no genera $\mathbb{R}^3$ — Una base para $\mathbb{R}^3$ — Genera $\mathbb{R}^3$, pero es linealmente dependiente

■

PROBLEMAS DE PRÁCTICA

1. Sean $\mathbf{v}_1 = \begin{bmatrix}1\\-2\\3\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}-2\\7\\-9\end{bmatrix}$. Determine si $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para $\mathbb{R}^3$. ¿Es $\{\mathbf{v}_1, \mathbf{v}_2\}$ una base para $\mathbb{R}^2$?

2. Sean $\mathbf{v}_1 = \begin{bmatrix}1\\-3\\4\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}6\\2\\-1\end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix}2\\-2\\3\end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix}-4\\-8\\9\end{bmatrix}$. Encuentre una base para el subespacio W generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

3. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, y $H = \left\{ \begin{bmatrix} s \\ s \\ 0 \end{bmatrix} : s \text{ en } \mathbb{R} \right\}$. Entonces cualquier vector en H es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ porque

$$\begin{bmatrix} s \\ s \\ 0 \end{bmatrix} = s \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

¿Es $\{\mathbf{v}_1, \mathbf{v}_2\}$ una base para H?

SG Dominio de bases 4 a 10 (Mastering: Basis 4-10)

4.3 Ejercicios

Determine cuáles conjuntos de los ejercicios 1 a 8 son bases para $\mathbb{R}^3$. De los conjuntos que *no* sean bases, determine cuáles son linealmente independientes y cuáles generan $\mathbb{R}^3$. Justifique sus respuestas.

1. $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

2. $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

3. $\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}$

4. $\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \begin{bmatrix} -7 \\ 5 \\ 4 \end{bmatrix}$

5. $\begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 5 \end{bmatrix}$

6. $\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} -4 \\ -5 \\ 6 \end{bmatrix}$

7. $\begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ -1 \\ 5 \end{bmatrix}$

8. $\begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -2 \end{bmatrix}$

Encuentre bases para los espacios nulos de las matrices dadas en los ejercicios 9 y 10. Haga referencia a las notas que siguen al ejemplo 3 de la sección 4.2.

9. $\begin{bmatrix} 1 & 0 & -3 & 2 \\ 0 & 1 & -5 & 4 \\ 3 & -2 & 1 & -2 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 0 & -5 & 1 & 4 \\ -2 & 1 & 6 & -2 & -2 \\ 0 & 2 & -8 & 1 & 9 \end{bmatrix}$

11. Encuentre una base para el conjunto de vectores en $\mathbb{R}^3$ en el plano $x + 2y + z = 0$. [*Sugerencia:* Piense en la ecuación como un "sistema" de ecuaciones homogéneas.]

12. Encuentre una base para el conjunto de vectores en $\mathbb{R}^2$ que están sobre la línea $y = 5x$.

En los ejercicios 13 y 14, suponga que A es equivalente por filas a B. Encuentre bases para Nul A y Col A.

13. $A = \begin{bmatrix} -2 & 4 & -2 & -4 \\ 2 & -6 & -3 & 1 \\ -3 & 8 & 2 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 6 & 5 \\ 0 & 2 & 5 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 2 & -5 & 11 & -3 \\ 2 & 4 & -5 & 15 & 2 \\ 1 & 2 & 0 & 4 & 5 \\ 3 & 6 & -5 & 19 & -2 \end{bmatrix}$,

$B = \begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 5 & -7 & 8 \\ 0 & 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 15 a 18, encuentre una base para el espacio generado por los vectores dados, $\mathbf{v}_1, \ldots, \mathbf{v}_5$.

15. $\begin{bmatrix} 1 \\ 0 \\ -3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} -3 \\ -4 \\ 1 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ -8 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -6 \\ 9 \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -3 \\ 3 \\ -4 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \\ 1 \end{bmatrix}$

17. [M] $\begin{bmatrix} 8 \\ 9 \\ -3 \\ -6 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 1 \\ -4 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ -9 \\ 6 \\ -7 \end{bmatrix}, \begin{bmatrix} 6 \\ 8 \\ 4 \\ -7 \\ 10 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 11 \\ -8 \\ -7 \end{bmatrix}$

18. **[M]** $\begin{bmatrix} -8 \\ 7 \\ 6 \\ 5 \\ -7 \end{bmatrix}, \begin{bmatrix} 8 \\ -7 \\ -9 \\ -5 \\ 7 \end{bmatrix}, \begin{bmatrix} -8 \\ 7 \\ 4 \\ 5 \\ -7 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 9 \\ 6 \\ -7 \end{bmatrix}, \begin{bmatrix} -9 \\ 3 \\ -4 \\ -1 \\ 0 \end{bmatrix}$

19. Sean $\mathbf{v}_1 = \begin{bmatrix} 4 \\ -3 \\ 7 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 9 \\ -2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 7 \\ 11 \\ 6 \end{bmatrix}$, y $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Puede verificarse que $4\mathbf{v}_1 + 5\mathbf{v}_2 - 3\mathbf{v}_3 = \mathbf{0}$. Use esta información para encontrar una base para H. Existe más de una respuesta.

20. Sean $\mathbf{v}_1 = \begin{bmatrix} 7 \\ 4 \\ -9 \\ -5 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 4 \\ -7 \\ 2 \\ 5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ -5 \\ 3 \\ 4 \end{bmatrix}$. Puede verificarse que $\mathbf{v}_1 - 3\mathbf{v}_2 + 5\mathbf{v}_3 = \mathbf{0}$. Use esta información y encuentre una base para $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

En los ejercicios 21 y 22, señale cada enunciado como verdadero o falso. Justifique sus respuestas.

21. a. Un solo vector, por sí mismo, es linealmente dependiente.

b. Si $H = \text{Gen}\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$, entonces $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ es una base para H.

c. Las columnas de una matriz invertible de $n \times n$ forman una base para $\mathbb{R}^n$.

d. Una base es un conjunto generador lo más grande posible.

e. En algunos casos, las relaciones de dependencia entre las columnas de una matriz pueden resultar afectadas por ciertas operaciones elementales de fila sobre la matriz.

22. a. Un conjunto linealmente independiente en un subespacio H es una base para H.

b. Si un conjunto finito S de vectores distintos de cero genera un espacio vectorial V, entonces algún subconjunto de S es una base para V.

c. Una base es un conjunto linealmente independiente lo más grande posible.

d. El método estándar para producir un conjunto generador para Nul A, descrito en la sección 4.2, algunas veces no logra producir una base para Nul A.

e. Si B es una forma escalonada de una matriz A, entonces las columnas pivote de B forman una base para Col A.

23. Suponga que $\mathbb{R}^4 = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$. Explique por qué $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es una base para $\mathbb{R}^4$.

24. Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto linealmente independiente en $\mathbb{R}^n$. Explique por qué $\mathcal{B}$ debe ser una base para $\mathbb{R}^n$.

25. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, y sea H el conjunto de vectores en $\mathbb{R}^3$ cuyas entradas segunda y tercera son iguales. Entonces todo vector de H tiene una ampliación única como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, porque

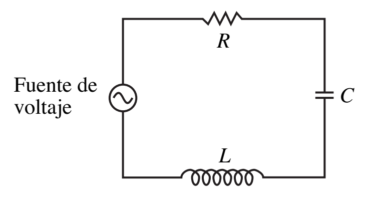
$$\begin{bmatrix} s \\ t \\ t \end{bmatrix} = s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (t - s) \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

para cualesquiera s y t. ¿Es $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ una base para H? ¿Por qué sí o por qué no?

26. En el espacio vectorial de todas las funciones con valores reales, encuentre una base para el subespacio generado por $\{\text{sen } t, \text{sen } 2t, \text{sen } t \cos t\}$.

27. Sea V el espacio vectorial de las funciones que describen la vibración de un sistema compuesto de masa-resorte. (Haga referencia al ejercicio 19 de la sección 4.1.) Encuentre una base para V.

28. (*Circuito RLC.*) El circuito de la figura consiste en un resistor (R ohms), un inductor (L henrys), un capacitor (C faradios), y una fuente de voltaje inicial. Sea $b = R/(2L)$, y suponga que R, L y C han sido elegidos de modo que b también sea igual a $1/\sqrt{LC}$. (Esto se hace, por ejemplo, cuando el circuito se usa en un voltímetro.) Sea $v(t)$ el voltaje (en volts) en el tiempo t, medido a lo largo del capacitor. Se puede mostrar que v está en el espacio nulo H de la transformación lineal que mapea $v(t)$ en $Lv''(t) + Rv'(t) + (1/C)v(t)$, y H consiste en todas las funciones de la forma $v(t) = e^{-bt}(c_1 + c_2 t)$. Encuentre una base para H.



Los ejercicios 29 y 30 muestran que toda base para $\mathbb{R}^n$ debe contener exactamente n vectores.

29. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto de k vectores en $\mathbb{R}^n$, con $k < n$. Use un teorema de la sección 1.4 para explicar por qué S no puede ser una base para $\mathbb{R}^n$.

30. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto de k vectores en $\mathbb{R}^n$, siendo $k > n$. Use un teorema del capítulo 1 para explicar por qué S no puede ser una base para $\mathbb{R}^n$.

Los ejercicios 31 y 32 revelan una conexión importante entre la independencia lineal y las transformaciones lineales, y permiten practicar el uso de la definición de dependencia lineal. Sean V y

W espacios vectoriales, sea $T : V \to W$ una transformación lineal, y sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ un subconjunto de V.

31. Muestre que si $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es linealmente dependiente en V, entonces el conjunto de imágenes $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$ es linealmente dependiente en W. Este hecho demuestra que si una transformación lineal mapea un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en un conjunto linealmente *independiente* $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$, entonces el conjunto original también es linealmente independiente (puesto que no puede ser linealmente dependiente).

32. Suponga que T es una transformación uno a uno, de modo que una ecuación $T(\mathbf{u}) = T(\mathbf{v})$ siempre implica que $\mathbf{u} = \mathbf{v}$. Muestre que si el conjunto de imágenes $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$ es linealmente dependiente, entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es linealmente dependiente. Este hecho muestra que *una transformación lineal uno a uno mapea un conjunto linealmente independiente sobre un conjunto linealmente independiente* (porque en este caso el conjunto de imágenes no puede ser linealmente dependiente.)

33. Considere los polinomios $\mathbf{p}_1(t) = 1 + t^2$ y $\mathbf{p}_2(t) = 1 - t^2$. ¿Es $\{\mathbf{p}_1, \mathbf{p}_2\}$ un conjunto linealmente independiente en $\mathbb{P}_3$? ¿Por qué sí o por qué no?

34. Considere los polinomios $\mathbf{p}_1(t) = 1 + t$, $\mathbf{p}_2(t) = 1 - t$, y $\mathbf{p}_3(t) = 2$ (para toda t). Por inspección, escriba una relación de dependencia lineal entre $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$. Después encuentre una base para Gen$\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$.

35. Sea V un espacio vectorial que contiene un conjunto linealmente independiente $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$. Describa cómo construir un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ en V tal que $\{\mathbf{v}_1, \mathbf{v}_3\}$ sea una base para Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

36. **[M]** Sea $H = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ y $K = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, donde

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 2 \\ 2 \\ 7 \\ -3 \end{bmatrix},$$

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 8 \\ -4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ -2 \\ 9 \\ -5 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 6 \\ -2 \end{bmatrix}$$

Encuentre bases para H, K y $H + K$. (Vea los ejercicios 33 y 34 de la sección 4.1.)

37. **[M]** Demuestre que $\{t, \text{sen } t, \cos 2t, \text{sen } t \cos t\}$ es un conjunto linealmente independiente de funciones definidas en $\mathbb{R}$. Comience por suponer que

$$c_1 \cdot t + c_2 \cdot \text{sen } t + c_3 \cdot \cos 2t + c_4 \cdot \text{sen } t \cos t = 0 \qquad (5)$$

La ecuación (5) debe ser válida para toda t real, así que elija varios valores específicos de t (por ejemplo, $t = 0, .1, .2$) hasta obtener un sistema con las suficientes ecuaciones como para determinar que todas las c_j deben ser cero.

38. **[M]** Muestre que $\{1, \cos t, \cos^2 t, \ldots, \cos^6 t\}$ es un conjunto linealmente independiente de funciones definidas en $\mathbb{R}$. Use el método del ejercicio 37. (Este resultado se necesitará en el ejercicio 34 de la sección 4.5.)

CD Espacio columna y espacio nulo (Column Space and Null Space)

CD Una base para Col A (A Basis for Col A)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sea $A = [\mathbf{v}_1 \ \ \mathbf{v}_2]$. Mediante operaciones por fila se tiene que

$$A = \begin{bmatrix} 1 & -2 \\ -2 & 7 \\ 3 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$$

No toda fila de A contiene una posición pivote, así que las columnas de A no generan $\mathbb{R}^3$, por el teorema 4 de la sección 1.4. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ no es una base para $\mathbb{R}^3$. Como $\mathbf{v}_1$ y $\mathbf{v}_2$ no están en $\mathbb{R}^2$, no pueden ser una base para $\mathbb{R}^2$. Sin embargo, como resulta evidente que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes, constituyen una base para un subespacio de $\mathbb{R}^3$, a saber, Gen$\{\mathbf{v}_1, \mathbf{v}_2\}$.

2. Prepare una matriz A cuyo espacio columna sea el espacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, y luego reduzca por filas a A para encontrar sus columnas pivote.

$$A = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 20 & 4 & -20 \\ 0 & -25 & -5 & 25 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 5 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Las primeras dos columnas de A son las columnas pivote y, por lo tanto, forman una base para Col $A = W$. Entonces $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para W. Observe que la forma escalonada reducida de A no es necesaria para localizar las columnas pivote.

3. Ni $\mathbf{v}_1$ ni $\mathbf{v}_2$ están en H, así que $\{\mathbf{v}_1, \mathbf{v}_2\}$ no puede ser una base para H. De hecho, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para el *plano* de todos los vectores de la forma $(c_1, c_2, 0)$, pero H es sólo una *línea*.

4.4 SISTEMAS DE COORDENADAS

Una razón importante para especificar una base $\mathcal{B}$ para un espacio vectorial V es imponer un "sistema de coordenadas" sobre V. En esta sección se mostrará que si $\mathcal{B}$ contiene n vectores, entonces el sistema de coordenadas hará que V actúe como $\mathbb{R}^n$. Si V ya es $\mathbb{R}^n$, entonces $\mathcal{B}$ determinará un sistema de coordenadas que proporciona una nueva "vista" de V.

La existencia de sistemas de coordenadas se apoya en el siguiente resultado fundamental.

TEOREMA 7 **El teorema de representación única**

Sea $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Entonces, para cada $\mathbf{x}$ en V, existe un único conjunto de escalares $c_1, \dots, c_n$ tal que

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n \tag{1}$$

DEMOSTRACIÓN Dado que $\mathcal{B}$ genera V, existen escalares tales que (1) es válida. Suponga que $\mathbf{x}$ también tiene la representación

$$\mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_n\mathbf{b}_n$$

para escalares $d_1, \dots, d_n$. Entonces, restando, se tiene

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_n - d_n)\mathbf{b}_n \tag{2}$$

Puesto que $\mathcal{B}$ es linealmente independiente, los pesos en (2) tienen que ser todos cero. Esto es, $c_j = d_j$ para $1 \le j \le n$. ■

DEFINICIÓN

Suponga que el conjunto $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ es una base para V y que $\mathbf{x}$ está en V. Las **coordenadas de x relativas a la base $\mathcal{B}$** (o las **$\mathcal{B}$-coordenadas de x**) son los pesos $c_1, \dots, c_n$ tales que $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$.

Si $c_1, \dots, c_n$ son las $\mathcal{B}$-coordenadas de $\mathbf{x}$, entonces el vector en $\mathbb{R}^n$

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

es el **vector de coordenadas de x** (**relativas a** $\mathcal{B}$) o el **vector de** $\mathcal{B}$**-coordenadas de x**. La función $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es la **función de coordenadas** (**determinada por** $\mathcal{B}$).[1]

EJEMPLO 1 Considere una base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ para $\mathbb{R}^2$, donde $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Suponga que una **x** en $\mathbb{R}^2$ tiene el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Encuentre **x**.

Solución Las $\mathcal{B}$-coordenadas de **x** indican cómo construir **x** a partir de los vectores en $\mathcal{B}$. Esto es,

$$\mathbf{x} = (-2)\mathbf{b}_1 + 3\mathbf{b}_2 = (-2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

EJEMPLO 2 Las entradas del vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$ son las coordenadas de **x** relativas a la *base estándar* $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, puesto que

$$\begin{bmatrix} 1 \\ 6 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{e}_1 + 6 \cdot \mathbf{e}_2$$

Si $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, entonces $[\mathbf{x}]_{\mathcal{E}} = \mathbf{x}$.

Una interpretación gráfica de las coordenadas

Un sistema de coordenadas en un conjunto es una función uno a uno de los puntos del conjunto en $\mathbb{R}^n$. Por ejemplo, el papel normal para gráficas proporciona un sistema de coordenadas para el plano cuando se elijen ejes perpendiculares y una unidad de medida en cada eje. En la figura 1 se muestra la base estándar $\{\mathbf{e}_1, \mathbf{e}_2\}$, los vectores $\mathbf{b}_1 (= \mathbf{e}_1)$ y $\mathbf{b}_2$ del ejemplo 1, y el vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$. Las coordenadas 1 y 6 proporcionan la ubicación de **x** relativa a la base estándar: 1 unidad en la dirección de $\mathbf{e}_1$ y 6 unidades en la dirección de $\mathbf{e}_2$.

En la figura 2 se muestran los vectores $\mathbf{b}_1$, $\mathbf{b}_2$ y **x** de la figura 1. (Geométricamente, los tres vectores pertenecen a una línea vertical en ambas figuras.) Sin embargo, se borró la cuadrícula de las coordenadas estándar y se sustituyó por una retícula especialmente adaptada a la base $\mathcal{B}$ en el ejemplo 1. El vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ proporciona la ubicación de **x** en este nuevo sistema de coordenadas: -2 unidades en la dirección $\mathbf{b}_1$ y 3 unidades en la dirección $\mathbf{b}_2$.

[1] El concepto de función de coordenadas supone que la base $\mathcal{B}$ es un conjunto indexado cuyos vectores se enlistan en algún orden fijo preasignado. Esta propiedad permite que la definición de $[\mathbf{x}]_{\mathcal{B}}$ no resulte ambigua.

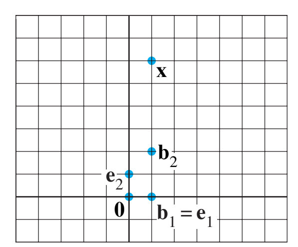


FIGURA 1 Papel para gráficas estándar.

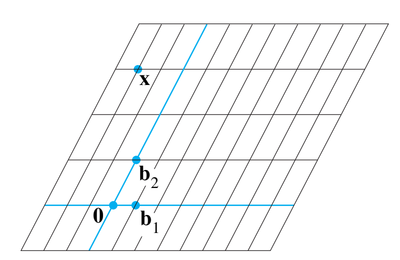


FIGURA 2 Papel para gráficas $\mathcal{B}$.

EJEMPLO 3 En cristalografía, la descripción de una red cristalina es mejorada al seleccionar una base $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ para $\mathbb{R}^3$ que corresponda a tres aristas adyacentes de una "celda unitaria" del cristal. Una red completa se construye al apilar varias copias de una celda unitaria. Hay catorce tipos básicos de celdas unitarias; en la figura 3 se muestran tres.[2]

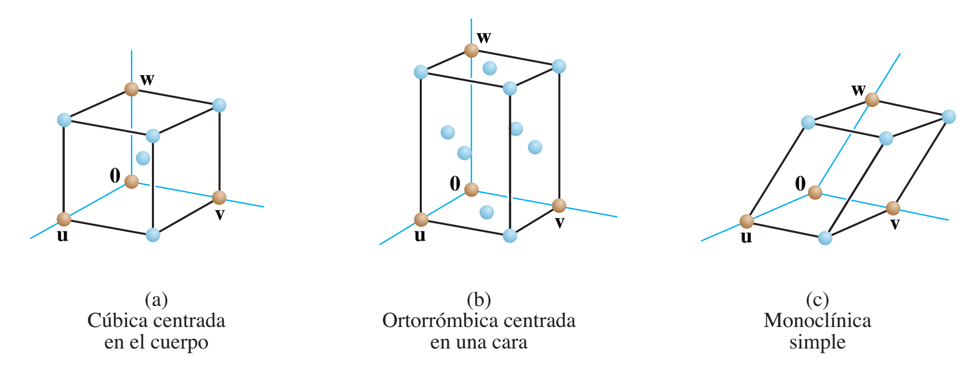


FIGURA 3 Ejemplos de celdas unitarias.

Dentro del cristal, las coordenadas de los átomos se dan en relación con la base para la red. Por ejemplo,

$$\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

identifica el átomo superior centrado en la cara de la celda mostrada en la figura 3(b).

Coordenadas en $\mathbb{R}^n$

Cuando una base $\mathcal{B}$ para $\mathbb{R}^n$ está fija, es fácil encontrar el vector de $\mathcal{B}$-coordenadas de una $\mathbf{x}$ específica, como en el ejemplo siguiente.

[2]Vea *The Science and Engineering of Materials*, 4a. ed., por Donald R. Askeland (Boston: Prindle, Weber & Schmidt, 2002), pág. 36.

EJEMPLO 4 Sea $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Encuentre el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ de $\mathbf{x}$ relativo a $\mathcal{B}$.

Solución Las $\mathcal{B}$-coordenadas c_1, c_2 de $\mathbf{x}$ satisfacen

$$c_1 \underset{\mathbf{b}_1}{\begin{bmatrix} 2 \\ 1 \end{bmatrix}} + c_2 \underset{\mathbf{b}_2}{\begin{bmatrix} -1 \\ 1 \end{bmatrix}} = \underset{\mathbf{x}}{\begin{bmatrix} 4 \\ 5 \end{bmatrix}}$$

o bien

$$\begin{bmatrix} \underset{\mathbf{b}_1}{\begin{matrix} 2 \\ 1 \end{matrix}} & \underset{\mathbf{b}_2}{\begin{matrix} -1 \\ 1 \end{matrix}} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \underset{\mathbf{x}}{\begin{bmatrix} 4 \\ 5 \end{bmatrix}} \tag{3}$$

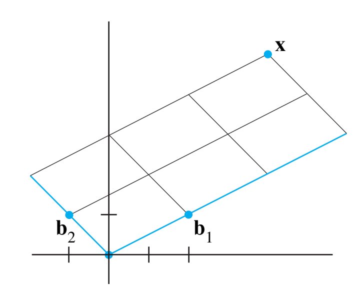


FIGURA 4
El vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$ es (3, 2).

Esta ecuación puede resolverse mediante operaciones por fila con una matriz aumentada o aplicando la inversa de la matriz a la izquierda. En cualquier caso, la solución es $c_1 = 3$, $c_2 = 2$. Entonces $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$, y

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Vea la figura 4.

La matriz en (3) cambia las $\mathcal{B}$-coordenadas de un vector $\mathbf{x}$ a las coordenadas estándar para $\mathbf{x}$. Puede realizarse un cambio análogo de coordenadas en $\mathbb{R}^n$ para una base $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$. Sea

$$P_{\mathcal{B}} = [\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n\,]$$

Entonces la ecuación vectorial

$$\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_n\mathbf{b}_n$$

es equivalente a

$$\boxed{\mathbf{x} = P_{\mathcal{B}}\,[\,\mathbf{x}\,]_{\mathcal{B}}} \tag{4}$$

$P_{\mathcal{B}}$ se denomina **matriz de cambio de coordenadas** de $\mathcal{B}$ a la base estándar en $\mathbb{R}^n$. La multiplicación por la izquierda por $P_{\mathcal{B}}$ transforma el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ en $\mathbf{x}$. La ecuación de cambio de coordenadas (4) es importante y será necesario aplicarla en varios puntos de los capítulos 5 y 7.

Como las columnas de $P_{\mathcal{B}}$ forman una base para $\mathbb{R}^n$, $P_{\mathcal{B}}$ es invertible (según el teorema de la matriz invertible). La multiplicación por la izquierda por $P_{\mathcal{B}}^{-1}$ convierte a $\mathbf{x}$ en su vector de $\mathcal{B}$-coordenadas:

$$P_{\mathcal{B}}^{-1}\mathbf{x} = [\,\mathbf{x}\,]_{\mathcal{B}}$$

La correspondencia $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$, producida aquí por $P_{\mathcal{B}}^{-1}$, es la función de coordenadas mencionada con anterioridad. Como $P_{\mathcal{B}}^{-1}$ es una matriz invertible, la función de coordenadas es una transformación lineal uno a uno de $\mathbb{R}^n$ sobre $\mathbb{R}^n$, por el teorema de la matriz

invertible. (Vea también el teorema 12 de la sección 1.9.) Más adelante se verá que esta propiedad de la función de coordenadas también es cierta para un espacio vectorial general que tenga una base.

La función de coordenadas

La selección de una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ para un espacio vectorial V introduce un sistema de coordenadas en V. La función de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ conecta el posiblemente desconocido espacio V con el conocido espacio $\mathbb{R}^n$. Vea la figura 5. Los puntos en V pueden identificarse ahora por sus nuevos "nombres".

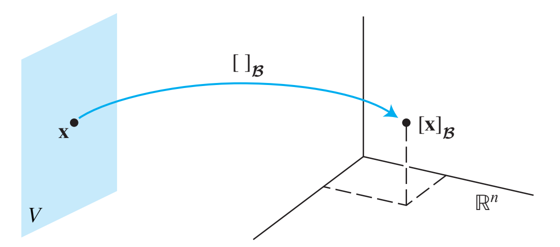


FIGURA 5 La función de coordenadas de V sobre $\mathbb{R}n$.

TEOREMA 8 Sea $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Entonces la función de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es una transformación lineal uno a uno de V sobre $\mathbb{R}^n$.

DEMOSTRACIÓN Tome dos vectores típicos de V, por ejemplo

$$\mathbf{u} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$
$$\mathbf{w} = d_1\mathbf{b}_1 + \cdots + d_n\mathbf{b}_n$$

Entonces, al usar operaciones vectoriales,

$$\mathbf{u} + \mathbf{w} = (c_1 + d_1)\mathbf{b}_1 + \cdots + (c_n + d_n)\mathbf{b}_n$$

Se deduce que

$$[\,\mathbf{u} + \mathbf{w}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [\,\mathbf{u}\,]_{\mathcal{B}} + [\,\mathbf{w}\,]_{\mathcal{B}}$$

De modo que la función de coordenadas conserva la suma. Si r es cualquier escalar, entonces

$$r\mathbf{u} = r(c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n) = (rc_1)\mathbf{b}_1 + \cdots + (rc_n)\mathbf{b}_n$$

Por lo tanto,

$$[\,r\mathbf{u}\,]_{\mathcal{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r\,[\,\mathbf{u}\,]_{\mathcal{B}}$$

Así que la función de coordenadas también conserva la multiplicación por escalares y, por lo tanto, es una transformación lineal. Vea en los ejercicios 23 y 24 la comprobación de que la función de coordenadas es uno a uno y mapea V sobre $\mathbb{R}^n$. ■

La linealidad de la función de coordenadas se amplía a combinaciones lineales, igual que en la sección 1.8. Si $\mathbf{u}_1, \ldots, \mathbf{u}_p$ están en V, y si $c_1, \ldots, c_p$ son escalares, entonces

$$[\, c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p \,]_{\mathcal{B}} = c_1 [\, \mathbf{u}_1 \,]_{\mathcal{B}} + \cdots + c_p [\, \mathbf{u}_p \,]_{\mathcal{B}} \tag{5}$$

Expresado con palabras, (5) indica que el vector de $\mathcal{B}$-coordenadas de una combinación lineal de $\mathbf{u}_1, \ldots, \mathbf{u}_p$ es la *misma* combinación lineal de sus vectores de coordenadas.

La función de coordenadas del teorema 8 es un importante ejemplo de *isomorfismo* de V sobre $\mathbb{R}^n$. En general, una transformación lineal uno a uno de un espacio vectorial V sobre un espacio vectorial W se denomina **isomorfismo** de V sobre W (*iso* viene del griego y significa "lo mismo", y *morfos* es la palabra griega para "forma" o "estructura"). La notación y terminología para V y W difieren, pero los dos espacios son indistinguibles como espacios vectoriales. *Todo cálculo de espacios vectoriales en V se reproduce exactamente en W, y viceversa.* Vea los ejercicios 25 y 26.

SG Espacios vectoriales isomorfos 4 a 12 (Isomorphic Vector Spaces 4-12)

EJEMPLO 5 Sea $\mathcal{B}$ la base estándar del espacio $\mathbb{P}_3$ de los polinomios; esto es, sea $\mathcal{B} = \{1, t, t^2, t^3\}$. Un elemento típico $\mathbf{p}$ de $\mathbb{P}_3$ tiene la forma

$$\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

Dado que $\mathbf{p}$ ya se muestra como una combinación lineal de los vectores de la base estándar, se concluye que

$$[\, \mathbf{p} \,]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Entonces la función de coordenadas $\mathbf{p} \mapsto [\mathbf{p}]_{\mathcal{B}}$ es un isomorfismo de $\mathbb{P}_3$ sobre $\mathbb{R}^4$. Todas las operaciones de espacio vectorial en $\mathbb{P}_3$ corresponden a operaciones en $\mathbb{R}^4$. ■

Al pensar en $\mathbb{P}_3$ y $\mathbb{R}^4$ como imágenes desplegadas en distintas pantallas de computadora que están conectadas por medio de la función de coordenadas, entonces cada operación de espacio vectorial en $\mathbb{P}_3$ en una de las pantallas se duplica de manera exacta mediante una operación de vectores correspondiente en $\mathbb{R}^4$ en la otra pantalla. Los vectores de la pantalla $\mathbb{P}_3$ lucen diferentes a los de la pantalla $\mathbb{R}^4$, pero "actúan" como vectores exactamente en la misma forma. Vea la figura 6.

EJEMPLO 6 Use vectores de coordenadas para comprobar que los polinomios $1 + 2t^2$, $4 + t + 5t^2$, y $3 + 2t$ son linealmente dependientes en $\mathbb{P}_2$.

Solución La función de coordenadas del ejemplo 5 produce los vectores de coordenadas (1, 0, 2), (4, 1, 5) y (3, 2, 0), respectivamente. Si estos vectores se escriben como las *columnas* de una matriz A, es posible determinar su independencia mediante la reduc-

FIGURA 6 El espacio $\mathbb{P}_3$ es isomorfo a $\mathbb{R}^4$.

ción por filas de la matriz aumentada para $A\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 5 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Las columnas de A son linealmente dependientes, así que los polinomios correspondientes son linealmente dependientes. De hecho, es fácil comprobar que la columna 3 de A es dos veces la columna 2 menos cinco veces la columna 1. La relación correspondiente para los polinomios es

$$3 + 2t = 2(4 + t + 5t^2) - 5(1 + 2t^2)$$

El ejemplo final tiene que ver con un plano en $\mathbb{R}^3$ que es isomorfo a $\mathbb{R}^2$.

EJEMPLO 7 Sea $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$, y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Entonces $\mathcal{B}$ es una base para $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$. Determine si $\mathbf{x}$ está en H y, si así fuera, encuentre el vector de coordenadas de $\mathbf{x}$ relativo a $\mathcal{B}$.

Solución Si $\mathbf{x}$ está en H, entonces la siguiente ecuación vectorial es consistente:

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

Los escalares c_1 y c_2, si existen, son las $\mathcal{B}$-coordenadas de $\mathbf{x}$. Mediante operaciones por fila, se obtiene

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Entonces $c_1 = 2$, $c_2 = 3$, y $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. En la figura 7, se muestra el sistema de coordenadas en H determinado por $\mathcal{B}$.

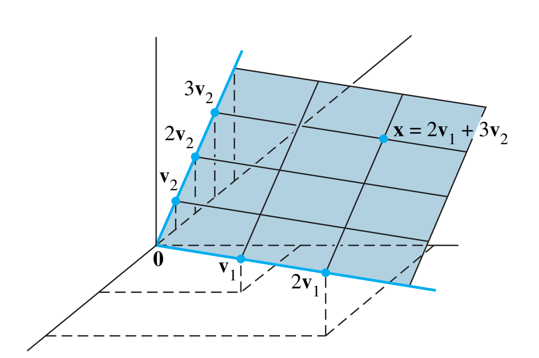


FIGURA 7 Un sistema de coordenadas en un plano H en $\mathbb{R}^3$.

Si se eligiera una base distinta para H, ¿el sistema de coordenadas asociado también haría a H isomorfo a $\mathbb{R}^2$? Seguramente esto es cierto, y se demostrará en la siguiente sección.

PROBLEMAS DE PRÁCTICA

1. Sean $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}$, y $\mathbf{x} = \begin{bmatrix} -8 \\ 2 \\ 3 \end{bmatrix}$.
   a. Demuestre que el conjunto $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ es una base de $\mathbb{R}^3$.
   b. Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a la base estándar.
   c. Escriba la ecuación que relaciona $\mathbf{x}$ en $\mathbb{R}^3$ con $[\mathbf{x}]_{\mathcal{B}}$.
   d. Encuentre $[\mathbf{x}]_{\mathcal{B}}$, para la $\mathbf{x}$ dada arriba.
2. El conjunto $\mathcal{B} = \{1 + t,\ 1 + t^2,\ t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas de $\mathbf{p}(t) = 6 + 3t - t^2$ relativo a $\mathcal{B}$.

4.4 EJERCICIOS

En los ejercicios 1 a 4, encuentre un vector $\mathbf{x}$ determinado por el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ y la base $\mathcal{B}$ dados.

1. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

2. $\mathcal{B} = \left\{ \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 8 \\ -5 \end{bmatrix}$

3. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -7 \\ 0 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$

4. $\mathcal{B} = \left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -7 \\ 3 \end{bmatrix} \right\}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -4 \\ 8 \\ -7 \end{bmatrix}$

En los ejercicios 5 a 8, encuentre el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ de $\mathbf{x}$ relativo a la base dada $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$.

5. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ -5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

6. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 5 \\ -6 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

7. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 9 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 8 \\ -9 \\ 6 \end{bmatrix}$

8. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix}$

En los ejercicios 9 y 10, encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a la base estándar en $\mathbb{R}^n$.

9. $\mathcal{B} = \left\{ \begin{bmatrix} 2 \\ -9 \end{bmatrix}, \begin{bmatrix} 1 \\ 8 \end{bmatrix} \right\}$

10. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}, \begin{bmatrix} 8 \\ -2 \\ 7 \end{bmatrix} \right\}$

En los ejercicios 11 y 12, use una matriz inversa para encontrar $[\mathbf{x}]_{\mathcal{B}}$ para las $\mathbf{x}$ y $\mathcal{B}$ dadas.

11. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ -6 \end{bmatrix}$

12. $\mathcal{B} = \left\{ \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

13. El conjunto $\mathcal{B} = \{1 + t^2, t + t^2, 1 + 2t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas de $\mathbf{p}(t) = 1 + 4t + 7t^2$ relativo a $\mathcal{B}$.

14. El conjunto $\mathcal{B} = \{1 - t^2, t - t^2, 2 - 2t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas $\mathbf{p}(t) = 3 + t - 6t^2$ relativo a $\mathcal{B}$.

En los ejercicios 15 y 16, señale cada enunciado como verdadero o falso. Justifique sus respuestas. A menos que se especifique lo contrario, $\mathcal{B}$ es una base para un espacio vectorial V.

15. a. Si $\mathbf{x}$ está en V y si $\mathcal{B}$ contiene n vectores, entonces el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$ está en $\mathbb{R}^n$.

b. Si $P_{\mathcal{B}}$ es la matriz de cambio de coordenadas, entonces $[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}\mathbf{x}$, para $\mathbf{x}$ en V.

c. Los espacios vectoriales $\mathbb{P}_3$, y $\mathbb{R}^3$ son isomorfos.

16. a. Si $\mathcal{B}$ es la base estándar para $\mathbb{R}^n$, entonces la $\mathcal{B}$-coordenada de una $\mathbf{x}$ en $\mathbb{R}^n$ es la propia $\mathbf{x}$.

b. La correspondencia $[\mathbf{x}]_{\mathcal{B}} \mapsto \mathbf{x}$ se llama función de coordenadas.

c. En algunos casos, un plano en $\mathbb{R}^3$ puede ser isomorfo a $\mathbb{R}^2$.

17. Los vectores $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ -8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$ generan $\mathbb{R}^2$, pero no forman una base. Encuentre dos modos diferentes de expresar $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ como una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

18. Sea $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Explique por qué los vectores de $\mathcal{B}$-coordenadas de $\mathbf{b}_1, \ldots, \mathbf{b}_n$ son las columnas $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de la matriz identidad de $n \times n$.

19. Sea S un conjunto finito en un espacio vectorial V con la propiedad de que toda $\mathbf{x}$ en V tiene una representación única como combinación lineal de elementos de S. Muestre que S es una base para V.

20. Suponga que $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es un conjunto generador linealmente dependiente para un espacio vectorial V. Muestre que cada $\mathbf{w}$ en V puede expresarse en más de una forma como combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_4$. [*Sugerencia:* Sea $\mathbf{w} = k_1\mathbf{v}_1 + \cdots + k_4\mathbf{v}_4$ un vector arbitrario en V. Use la independencia lineal de $\{\mathbf{v}_1, ... , \mathbf{v}_4\}$ para producir otra representación de $\mathbf{w}$ como una combinación lineal de $\mathbf{v}_1 \ldots, \mathbf{v}_4$.]

21. Sea $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \end{bmatrix} \right\}$. Como la función de coordenadas determinada por $\mathcal{B}$ es una transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^2$, esta función tiene que implementarse mediante alguna matriz A de 2×2. Encuentre dicha matriz. [*Sugerencia:* La multiplicación por A deberá transformar un vector $\mathbf{x}$ en su vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$.]

22. Sea $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ una base para $\mathbb{R}^n$. Forme una matriz A de $n \times n$ que implemente la función de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$. (Vea el ejercicio 21.)

Los ejercicios 23 a 26 se refieren a un espacio vectorial V, a una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, y a la función de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$.

23. Muestre que la función de coordenadas es uno a uno. [*Sugerencia:* Suponga que $[\mathbf{u}]_{\mathcal{B}} = [\mathbf{w}]_{\mathcal{B}}$ para algunas $\mathbf{u}$ y $\mathbf{w}$ en V, y muestre que $\mathbf{u} = \mathbf{w}$.)

24. Muestre que la función de coordenadas es *sobre* $\mathbb{R}^n$. Esto es, dada cualquier $\mathbf{y}$ en $\mathbb{R}^n$, con entradas $y_1, \ldots, y_n$, encuentre $\mathbf{u}$ en V tal que $[\mathbf{u}]_{\mathcal{B}} = \mathbf{y}$.

25. Muestre que un subconjunto $\{\mathbf{u}_1, \ldots, \mathbf{u}_p)$ en V es linealmente independiente si, y sólo si, el conjunto de vectores de $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_p]_{\mathcal{B}}\}$ es linealmente independiente en $\mathbb{R}^n$. *Indicación:* Dado que la función de coordenadas es uno a uno, las siguien-

tes ecuaciones tienen las mismas soluciones, $c_1, \ldots, c_p$.

$$c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = \mathbf{0} \qquad \text{El vector cero en } V$$
$$[c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p]_{\mathcal{B}} = [\mathbf{0}]_{\mathcal{B}} \qquad \text{El vector cero en } \mathbb{R}^n$$

26. Dados los vectores $\mathbf{u}_1, \ldots, \mathbf{u}_p$, y $\mathbf{w}$ en V, muestre que $\mathbf{w}$ es una combinación lineal de $\mathbf{u}_1, \ldots, \mathbf{u}_p$ si, y sólo si, $[\mathbf{w}]_{\mathcal{B}}$ es una combinación lineal de los vectores de coordenadas $[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_p]_{\mathcal{B}}$.

En los ejercicios 27 a 30, use vectores de coordenadas para verificar la independencia lineal de los siguientes conjuntos de polinomios. Explique las operaciones realizadas.

27. $1 + t^3, 3 + t - 2t^2, -t + 3t^2 - t^3$

28. $1 - 2t^2 - 3t^3, t + t^3, 1 + 3t - 2t^2$

29. $(t - 1)^2, t^3 - 2, (t - 2)^3$

30. $(1 - t)^3, (2 - 3t)^2, 3t^2 - 4t^3$

31. Use vectores de coordenadas para verificar si los siguientes conjuntos de polinomios generan $\mathbb{P}_2$. Justifique sus conclusiones.

a. $1 - 3t + 5t^2, -3 + 5t - 7t^2, -4 + 5t - 6t^2, 1 - t^2$

b. $5t + t^2, 1 - 8t - 2t^2, -3 + 4t + 2t^2, 2 - 3t$

32. Sean $\mathbf{p}_1(t) = 1 + t^2$, $\mathbf{p}_2(t) = 2 - t + 3t^2$, $\mathbf{p}_3(t) = 1 + 2t - 4t^2$.

a. Use vectores de coordenadas para mostrar que estos polinomios forman una base para $\mathbb{P}_2$.

b. Considere la base $\mathcal{B} = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ para $\mathbb{P}_2$. Encuentre $\mathbf{q}$ en $\mathbb{P}_2$, dado que $[\mathbf{q}]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}$.

En los ejercicios 33 y 34 determine si los conjuntos de polinomios forman una base para $\mathbb{P}_3$. Justifique sus conclusiones.

33. **[M]** $3 + 7t, 5 + t - 2t^3, t - 2t^2, 1 + 16t - 6t^2 + 2t^3$

34. **[M]** $5 - 3t + 4t^2 + 2t^3, 9 + t + 8t^2 - 6t^3, 6 - 2t + 5t^2, t^3$

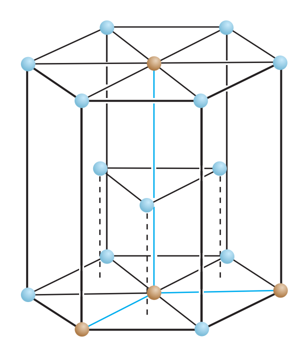
35. **[M]** Sean $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Muestre que $\mathbf{x}$ está en H y encuentre el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$ para

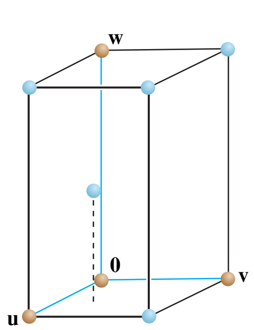
$$\mathbf{v}_1 = \begin{bmatrix} 11 \\ -5 \\ 10 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 14 \\ -8 \\ 13 \\ 10 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 19 \\ -13 \\ 18 \\ 15 \end{bmatrix}$$

36. **[M]** Sean $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Muestre que $\mathcal{B}$ es una base para H y que $\mathbf{x}$ está en H, y encuentre el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$ para

$$\mathbf{v}_1 = \begin{bmatrix} -6 \\ 4 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ -3 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 5 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 7 \\ -8 \\ 3 \end{bmatrix}$$

[M] Los ejercicios 37 y 38 se relacionan con la red cristalina del titanio, la cual tiene la estructura hexagonal mostrada a la izquierda de la figura acompañante. Los vectores $\begin{bmatrix} 2.6 \\ -1.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 4.8 \end{bmatrix}$ en $\mathbb{R}^3$ forman una base para la celda unitaria que se muestra a la derecha. Los números están dados en Ångstrom (1 Å $= 10^{-8}$ cm). En aleaciones de titanio, puede haber algunos átomos adicionales en la celda unitaria en los sitios *octaédricos* y *tetraédricos* (llamados así por los objetos geométricos que forman los átomos en esas ubicaciones).





La red hexagonal de empaque cerrado y su celda unitaria.

37. Uno de los sitios octaédricos es $\begin{bmatrix} 1/2 \\ 1/4 \\ 1/6 \end{bmatrix}$, con respecto a la base de la red. Determine las coordenadas de este sitio relativas a la base estándar de $\mathbb{R}^3$.

38. Uno de los sitios tetraédricos es $\begin{bmatrix} 1/2 \\ 1/2 \\ 1/3 \end{bmatrix}$. Determine las coordenadas de este sitio relativas a la base estándar de $\mathbb{R}^3$.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. a. Es evidente que la matriz $P_{\mathcal{B}} = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]$ es equivalente por filas a la matriz identidad. De acuerdo con el teorema de la matriz invertible, $P_{\mathcal{B}}$ es invertible y sus columnas forman una base para $\mathbb{R}^3$.

b. Del inciso (a), la matriz de cambio de coordenadas es $P_{\mathcal{B}} = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}$.

c. $\mathbf{x} = P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$

d. Para resolver la parte (c), probablemente sea más fácil reducir por filas una matriz aumentada en lugar de calcular $P_{\mathcal{B}}^{-1}$:

$$\underbrace{\begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}}_{P_{\mathcal{B}}} \underbrace{\begin{matrix} -8 \\ 2 \\ 3 \end{matrix}}_{\mathbf{x}} \sim \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{I} \underbrace{\begin{matrix} -5 \\ 2 \\ 1 \end{matrix}}_{[\mathbf{x}]_{\mathcal{B}}}$$

Por lo tanto,

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$$

2. Las coordenadas de $\mathbf{p}(t) = 6 + 3t - t^2$ con respecto de $\mathcal{B}$ satisfacen

$$c_1(1 + t) + c_2(1 + t^2) + c_3(t + t^2) = 6 + 3t - t^2$$

Al igualar los coeficientes de potencias iguales de t se obtiene

$$\begin{aligned} c_1 + c_2 &= 6 \\ c_1 + c_3 &= 3 \\ c_2 + c_3 &= -1 \end{aligned}$$

Al resolver, se encuentra que $c_1 = 5$, $c_2 = 1$, $c_3 = -2$, y $[\mathbf{p}]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$.

4.5 LA DIMENSIÓN DE UN ESPACIO VECTORIAL

El teorema 8 de la sección 4.4 implica que un espacio vectorial con una base $\mathcal{B}$ que contiene n vectores es isomorfo a $\mathbb{R}^n$. En esta sección se demostrará que este número n es una propiedad intrínseca (llamada dimensión) del espacio V que no depende de la base específica elegida. El análisis de la dimensión proporcionará una comprensión adicional de las propiedades de las bases.

El primer teorema generaliza un resultado muy bien conocido acerca del espacio vectorial $\mathbb{R}^n$.

TEOREMA 9 Si un espacio vectorial V tiene una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, entonces cualquier conjunto que contenga más de n vectores debe ser linealmente dependiente.

DEMOSTRACIÓN Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ un conjunto en V con más de n vectores. Los vectores de coordenadas $[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_p]_{\mathcal{B}}$ forman un conjunto linealmente dependiente en $\mathbb{R}^n$, puesto que hay más vectores (p) que entradas (n) en cada vector. Por lo tanto, existen

escalares $c_1, \ldots, c_p$, no todos cero, tales que

$$c_1 [\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_p [\mathbf{u}_p]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{El vector cero en } \mathbb{R}^n$$

Como la función de coordenadas es una transformación lineal,

$$[c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

El vector cero de la derecha contiene los n pesos necesarios para construir el vector $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ a partir de los vectores de la base en $\mathcal{B}$. Esto es, $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = 0 \cdot \mathbf{b}_1 + \cdots + 0 \cdot \mathbf{b}_n = \mathbf{0}$. Como las c_i no son todas cero, $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es linealmente dependiente.[1] ■

El teorema 9 implica que si un espacio vectorial V tiene una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, entonces cada conjunto linealmente independiente en V no tiene más de n vectores.

TEOREMA 10

Si un espacio vectorial V tiene una base con n vectores, entonces toda base de V debe consistir en exactamente n vectores.

DEMOSTRACIÓN Sean $\mathcal{B}_1$ una base con n vectores y $\mathcal{B}_2$ cualquier otra base (de V). Como $\mathcal{B}_1$ es una base y $\mathcal{B}_2$ es linealmente independiente, $\mathcal{B}_2$ no tiene más de n vectores, según el teorema 9. Además, puesto que $\mathcal{B}_2$ es una base y $\mathcal{B}_1$ es linealmente independiente, $\mathcal{B}_2$ tiene por lo menos n vectores. Así que $\mathcal{B}_2$ consiste en exactamente n vectores. ■

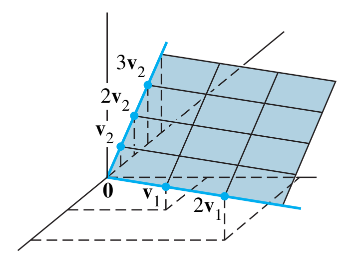
Si un espacio vectorial V distinto de cero es generado por un conjunto finito S, entonces un subconjunto de S es una base para V, de acuerdo con el teorema del conjunto generador. En este caso, el teorema 10 asegura que la siguiente definición tiene sentido.

DEFINICIÓN

Si V es generado por un conjunto finito, se dice que V es de **dimensión finita**, y la **dimensión** de V, que se escribe dim V, es el número de vectores en una base de V. La dimensión del espacio vectorial cero $\{\mathbf{0}\}$ se define como cero. Si V no es generado por un conjunto finito, entonces se dice que V es de **dimensión infinita**.

EJEMPLO 1 La base estándar para $\mathbb{R}^n$ contiene n vectores, entonces dim $\mathbb{R}^n = n$. La base polinomial estándar es $\{1, t, t^2\}$, lo cual muestra que dim $\mathbb{P}_2 = 3$. En general, dim $\mathbb{P}_n = n + 1$. El espacio $\mathbb{P}$ de todos los polinomios es de dimensión infinita (ejercicio 27). ■

[1] El teorema 9 también se aplica a conjuntos infinitos en V. Se dice que un conjunto infinito es linealmente dependiente si algún subconjunto finito es linealmente dependiente; en caso contrario, el conjunto es linealmente independiente. Si S es un conjunto infinito en V, tome cualquier subconjunto $(\mathbf{u}_1, \ldots, \mathbf{u}_p)$ de S, con °$p > n$. La demostración anterior establece que este subconjunto es linealmente dependiente y, por lo tanto, S también lo es.



EJEMPLO 2 Sea $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Entonces H es el plano estudiado en el ejemplo 7 de la sección 4.4. Una base para H es $\{\mathbf{v}_1, \mathbf{v}_2\}$, puesto que $\mathbf{v}_1$ y $\mathbf{v}_2$ no son múltiplos y, por lo tanto, son linealmente independientes. Entonces $\dim H = 2$.

EJEMPLO 3 Encuentre la dimensión del subespacio

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} : a, b, c, d \text{ en } \mathbb{R} \right\}$$

Solución Es fácil darse cuenta de que H es el conjunto de todas las combinaciones lineales de los vectores

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Es evidente, $\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{v}_2$ no es un múltiplo de $\mathbf{v}_1$, pero $\mathbf{v}_3$ es un múltiplo de $\mathbf{v}_2$. Según el teorema del conjunto generador, es posible desechar $\mathbf{v}_3$, y aún así tener un conjunto que genera H. Por último, $\mathbf{v}_4$ no es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$. Así que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ es linealmente independiente (por el teorema 4 de la sección 4.3) y, por lo tanto, es una base para H. Entonces $\dim H = 3$.

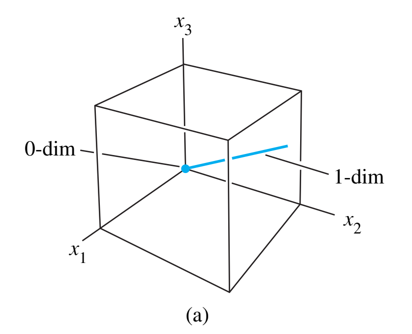
EJEMPLO 4 Los subespacios de $\mathbb{R}^3$ pueden clasificarse de acuerdo con su dimensión. Vea la figura 1.

Subespacios de dimensión 0. Sólo el subespacio cero.

Subespacios de dimensión 1. Cualquier subespacio generado por un único vector distinto de cero. Tales subespacios son líneas que pasan por el origen.

Subespacios de dimensión 2. Cualquier subespacio generado por dos vectores linealmente independientes. Tales subespacios son planos que pasan por el origen.

Subespacios de dimensión 3. Sólo el propio $\mathbb{R}^3$. Cualesquiera tres vectores linealmente independientes en $\mathbb{R}^3$ generan todo $\mathbb{R}^3$, de acuerdo con el teorema de la matriz invertible.



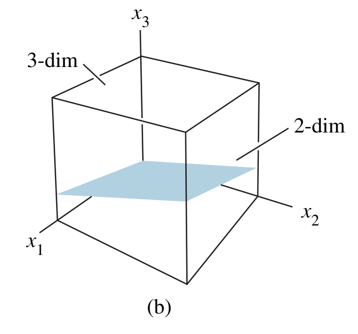


FIGURA 1 Subespacios representativos de $\mathbb{R}^3$.

Subespacios de un espacio de dimensión finita

El teorema siguiente es una contraparte natural del teorema del conjunto generador.

TEOREMA 11

Sea H un subespacio de un espacio vectorial de dimensión finita V. Cualquier conjunto linealmente independiente en H puede ampliarse, de ser necesario, hasta constituir una base para H. También, H es de dimensión finita y

$$\dim H \leq \dim V$$

DEMOSTRACIÓN Si $H = \{\mathbf{0}\}$ entonces, desde luego, $\dim H = 0 \leq \dim V$. En caso contrario, sea $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ cualquier conjunto linealmente independiente en H. Si S genera H, entonces S es una base para H. Si no, existe algún $\mathbf{u}_{k+1}$ en H que no está en Gen S. Pero entonces $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1})$ será linealmente independiente, porque ningún vector del conjunto puede ser una combinación lineal de vectores que le precedan (por el teorema 4).

Mientras el nuevo conjunto no genere H, es posible continuar con este proceso de ampliación de S hasta un conjunto linealmente independiente más grande en H. Pero el número de vectores en una ampliación linealmente independiente de S no puede exceder la dimensión de V, según el teorema 9. Entonces, en algún momento, una ampliación de S generará H y, por lo tanto, será una base para H, y $\dim H \leq \dim V$. ■

Cuando no se conoce la dimensión de un espacio o de un subespacio vectorial, la búsqueda de una base se simplifica con el teorema siguiente, el cual establece que si un conjunto tiene el número correcto de elementos, entonces es suficiente demostrar que el conjunto es linealmente independiente o bien que genera al espacio. El teorema resulta de importancia crítica en numerosos problemas de aplicaciones (relacionadas con ecuaciones diferenciales o en diferencias, por ejemplo) donde la independencia lineal es mucho más fácil de comprobar que la propiedad de generación.

TEOREMA 12

El teorema de la base

Sea V un espacio vectorial de dimensión p, $p \geq 1$. Cualquier conjunto linealmente independiente con exactamente p elementos en V es, de manera automática, una base para V. Cualquier conjunto de exactamente p elementos que genere V es automáticamente una base para V.

DEMOSTRACIÓN De acuerdo con el teorema 11, un conjunto linealmente independiente S de p elementos puede ampliarse hasta constituir una base para V. Pero esta base debe contener exactamente p elementos, puesto que $\dim V = p$. Por lo tanto, S tiene que ser ya la base para V. Suponga ahora que S tiene p elementos y genera V. Como V es distinto de cero, el teorema del conjunto generador implica la existencia de un subconjunto S' de S que es una base para V. Como $\dim V = p$, S' debe contener p vectores. Por lo tanto, $S = S'$. ■

Las dimensiones de Nul A y Col A

Dado que las columnas pivote de una matriz A forman una base para Col A, la dimensión de Col A se sabrá tan pronto se conozcan las columnas pivote. Podrá parecer que encontrar

la dimensión de Nul A requiere más trabajo, puesto que hallar una base para Nul A lleva, normalmente, más tiempo que encontrar una base para Col A. ¡Pero a continuación se presenta un atajo!

Sea A una matriz de $m \times n$ y suponga que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene k variables libres. Se sabe, por la sección 4.2, que el método estándar para encontrar un conjunto generador para Nul A produce exactamente k vectores linealmente independientes —por ejemplo, $\mathbf{u}_1, \dots, \mathbf{u}_k$—, uno por cada variable libre. Así, $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ es una base para Nul A, y el número de variables libres determina el tamaño de la base. Este resumen de hechos servirá como referencia en el futuro.

> La dimensión de Nul A es el número de variables libres incluidas en la ecuación $A\mathbf{x} = \mathbf{0}$, y la dimensión de Col A es el número de columnas pivote de A.

EJEMPLO 5 Encuentre las dimensiones del espacio nulo y del espacio columna de

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

Solución Reduzca por filas la matriz aumentada $[A \quad \mathbf{0}]$ a una forma escalonada:

$$\begin{bmatrix} 1 & -2 & 2 & 3 & -1 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Existen tres variables libres —x_2, x_4 y x_5—. Por lo tanto, la dimensión de Nul A es 3. También, dim Col $A = 2$ porque A tiene dos columnas pivote.

PROBLEMAS DE PRÁCTICA

Defina si los siguientes enunciados son verdaderos o falsos, y proporcione una razón para cada respuesta. Aquí V es un espacio vectorial distinto de cero y con dimensión finita.

1. Si dim $V = p$ y S es un subconjunto linealmente dependiente de V, entonces S contiene más de p vectores.
2. Si S genera V y T es un subconjunto de V que contiene más vectores que S, entonces T es linealmente dependiente.

4.5 EJERCICIOS

Para cada subespacio de los ejercicios 1 a 8, (a) encuentre una base, (b) establezca la dimensión.

1. $\left\{ \begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix} : s, t \text{ en } \mathbb{R} \right\}$

2. $\left\{ \begin{bmatrix} 4s \\ -3s \\ -t \end{bmatrix} : s, t \text{ en } \mathbb{R} \right\}$

3. $\left\{ \begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix} : a, b, c \text{ en } \mathbb{R} \right\}$

4. $\left\{ \begin{bmatrix} a+b \\ 2a \\ 3a-b \\ -b \end{bmatrix} : a, b \text{ en } \mathbb{R} \right\}$

5. $\left\{ \begin{bmatrix} a-4b-2c \\ 2a+5b-4c \\ -a+2c \\ -3a+7b+6c \end{bmatrix} : a, b, c \text{ en } \mathbb{R} \right\}$

6. $\left\{ \begin{bmatrix} 3a+6b-c \\ 6a-2b-2c \\ -9a+5b+3c \\ -3a+b+c \end{bmatrix} : a, b, c \text{ en } \mathbb{R} \right\}$

7. $\{(a, b, c) : a-3b+c=0, b-2c=0, 2b-c=0\}$

8. $\{(a, b, c, d) : a-3b+c=0\}$

9. Encuentre la dimensión del subespacio de todos los vectores en $\mathbb{R}^3$ cuyas entradas primera y tercera son iguales.

10. Encuentre la dimensión del subespacio H de $\mathbb{R}^2$ generado por

$$\begin{bmatrix} 2 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 10 \end{bmatrix}, \begin{bmatrix} -3 \\ 6 \end{bmatrix}.$$

En los ejercicios 11 y 12, encuentre la dimensión del subespacio generado por los vectores dados.

11. $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} -7 \\ -3 \\ 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} -8 \\ 6 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 7 \end{bmatrix}$

Determine las dimensiones de Nul A y Col A para las matrices que se muestran en los ejercicios 13 a 18.

13. $A = \begin{bmatrix} 1 & -6 & 9 & 0 & -2 \\ 0 & 1 & 2 & -4 & 5 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 3 & -4 & 2 & -1 & 6 \\ 0 & 0 & 1 & -3 & 7 & 0 \\ 0 & 0 & 0 & 1 & 4 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

15. $A = \begin{bmatrix} 1 & 0 & 9 & 5 \\ 0 & 0 & 1 & -4 \end{bmatrix}$

16. $A = \begin{bmatrix} 3 & 4 \\ -6 & 10 \end{bmatrix}$

17. $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$ **18.** $A = \begin{bmatrix} 1 & 4 & -1 \\ 0 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 19 y 20, V es un espacio vectorial. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

19. a. El número de columnas pivote de una matriz es igual a la dimensión de su espacio columna.

b. Un plano en $\mathbb{R}^3$ es un subespacio de dos dimensiones de $\mathbb{R}^3$.

c. La dimensión del espacio vectorial $\mathbb{P}_4$ es 4.

d. Si dim $V = n$ y S es un conjunto linealmente independiente en V, entonces S es una base para V.

e. Si un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ genera un espacio vectorial de dimensión finita V, y si T es un conjunto con más de p vectores en V, entonces T es linealmente dependiente.

20. a. $\mathbb{R}^2$ es un subespacio de dos dimensiones de $\mathbb{R}^3$.

b. El número de variables incluidas en la ecuación $A\mathbf{x} = \mathbf{0}$ es igual a la dimensión de Nul A.

c. Un espacio vectorial es de dimensión infinita si es generado por un conjunto infinito.

d. Si dim $V = n$ y S genera V, entonces S es una base para V.

e. El único subespacio tridimensional de $\mathbb{R}^3$ es el propio $\mathbb{R}^3$.

21. Los primeros cuatro polinomios de Hermite son 1, $2t$, $-2 + 4t^2$, y $-12t + 8t^3$. Estos polinomios surgen de manera natural al estudiar ciertas ecuaciones diferenciales importantes de la física matemática.[2] Muestre que los primeros cuatro polinomios de Hermite forman una base de $\mathbb{P}_3$.

22. Los primeros cuatro polinomios de Laguerre son 1, $1 - t$, $2 - 4t + t^2$, y $6 - 18t + 9t^2 - t^3$. Muestre que estos polinomios forman una base de $\mathbb{P}_3$.

23. Sea $\mathcal{B}$ la base de $\mathbb{P}_3$ que consta de los polinomios de Hermite del ejercicio 21, y sea $\mathbf{p}(t) = 7 - 12t - 8t^2 + 12t^3$. Encuentre el vector de coordenadas de $\mathbf{p}$ relativo a $\mathcal{B}$.

24. Sea $\mathcal{B}$ la base de $\mathbb{P}_2$ que consta de los tres primeros polinomios de Laguerre del ejercicio 22, y sea $\mathbf{p}(t) = 7 - 8t + 3t^2$. Encuentre el vector de coordenadas de $\mathbf{p}$ relativo a $\mathcal{B}$.

25. Sea S un subconjunto de un espacio vectorial V de dimensión n, y suponga que S contiene menos de n vectores. Explique por qué S no puede generar V.

26. Sea H un subespacio de dimensión n de un espacio vectorial V de dimensión n. Demuestre que $H = V$.

27. Explique por qué el espacio $\mathbb{P}$ de todos los polinomios es un espacio de dimensión infinita.

[2] Vea *Introduction to Functional Analysis*, 2a. ed., por A. E. Taylor y David C. Lay (Nueva York: John Wiley & Sons, 1980), págs. 92-93. También se tratan otros conjuntos de polinomios.

28. Muestre que el espacio $C(\mathbb{R})$ de todas las funciones continuas definidas en la recta real es un espacio de dimensión infinita.

En los ejercicios 29 y 30, V es un espacio vectorial distinto de cero de dimensión finita, y los vectores que se dan pertenecen a V. Señale cada enunciado como verdadero o falso. Justifique sus respuestas. (Estas preguntas son más difíciles que las de los ejercicios 19 y 20.)

29. a. Si existe un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}p\}$ que genera V, entonces $\dim V \leq p$.

b. Si existe un conjunto linealmente independiente $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en V, entonces $\dim V \geq p$.

c. Si $\dim V = p$, entonces existe un conjunto generador con $p + 1$ vectores en V.

30. a. Si existe un conjunto linealmente dependiente $\{\mathrm{v}_1, \ldots, \mathrm{v}_p\}$ en V, entonces $\dim V \leq p$.

b. Si ningún conjunto de p elementos en V logra generar V, entonces $\dim V > p$.

c. Si $p \geq 2$ y $\dim V = p$, entonces todo conjunto de $p - 1$ vectores distintos de cero es linealmente independiente.

Los ejercicios 31 y 32 se refieren a espacios vectoriales V y W de dimensión finita y a una transformación lineal $T : V \to W$.

31. Sea H un subespacio de V distinto de cero, y sea $T(H)$ el conjunto de imágenes de los vectores de H. Entonces $T(H)$ es un subespacio de W, según el ejercicio 35 de la sección 4.2. Demuestre que $\dim T(H) \leq \dim H$.

32. Sea H un subespacio de V distinto de cero, y suponga que T es una transformación uno a uno (lineal) de V en W. Demuestre que $\dim T(H) = \dim H$. Si sucediera que T es una transformación uno a uno de V *sobre* W, entonces $\dim V = \dim W$. Los espacios vectoriales isomorfos de dimensión finita tienen la misma dimensión.

33. **[M]** De acuerdo con el teorema 11, un conjunto linealmente independiente $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ en $\mathbb{R}^n$ puede ampliarse hasta constituir una base para $\mathbb{R}^n$. Una forma de hacer esto es creando $A = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k \ \ \mathbf{e}_1 \ \cdots \ \mathbf{e}_n]$, con $\mathbf{e}_1, \ldots, \mathbf{e}_n$ como las columnas de la matriz identidad. Las columnas pivote de A forman una base para $\mathbb{R}^n$.

a. Use el método descrito para ampliar los siguientes vectores y formar una base para $\mathbb{R}^5$:

$$\mathbf{v}_1 = \begin{bmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{bmatrix}$$

b. Explique por qué funciona el método en general: ¿Por qué los vectores originales $\mathbf{v}_1, \ldots, \mathbf{v}_k$ están incluidos en la base encontrada para Col A? ¿Por qué Col $A = \mathbb{R}^n$?

34. **[M]** Sean $\mathcal{B} = \{1, \cos t, \cos^2 t, \ldots, \cos^6 t\}$ y $\mathcal{C} = \{1, \cos t, \cos 2t, \ldots, \cos 6t\}$. Suponga las siguientes identidades trigonométricas (vea el ejercicio 37 de la sección 4.1).

$\cos 2t = -1 + 2\cos^2 t$

$\cos 3t = -3\cos t + 4\cos^3 t$

$\cos 4t = 1 - 8\cos^2 t + 8\cos^4 t$

$\cos 5t = 5\cos t - 20\cos^3 t + 16\cos^5 t$

$\cos 6t = -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t$

Sea H el subespacio de funciones generado por las funciones en $\mathcal{B}$. Entonces $\mathcal{B}$ es una base para H, según el ejercicio 38 dado en la sección 4.3.

a. Escriba los vectores de $\mathcal{B}$-coordenadas de los vectores en $\mathcal{C}$, y úselos para demostrar que $\mathcal{C}$ es un conjunto linealmente independiente en H.

b. Explique por qué $\mathcal{C}$ es una base para H.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Falso. Considere el conjunto $\{\mathbf{0}\}$.
2. Verdadero. Según el teorema del conjunto generador, S contiene una base para V; a la cual se le llamará S'. Entonces T contendrá más vectores que S'. De acuerdo con el teorema 9, T es linealmente dependiente.

4.6 RANGO

Con ayuda de los conceptos de espacios vectoriales, en esta sección se estudiará el *interior* de una matriz para descubrir muchas relaciones útiles e interesantes que se hallan ocultas entre sus filas y columnas.

Por ejemplo, imagine que se colocan 2000 números aleatorios en una matriz A de 40×50, y después se determina tanto el número máximo de columnas linealmente inde-

pendientes en A como el número máximo de columnas linealmente independientes en A^T (filas en A). Resulta notable que ambos números coincidan. Como pronto se verá, este valor común es el *rango* de la matriz. Para explicar la razón de esto, es necesario examinar el subespacio generado por las filas de A.

El espacio fila

Si A es una matriz de $m \times n$, cada fila de A tiene n entradas y así puede identificarse con un vector en $\mathbb{R}^n$. El conjunto de todas las combinaciones lineales de los vectores fila se denomina **espacio fila** de A y se denota por Fil A. Cada fila tiene n entradas, así que Fil A es un subespacio de $\mathbb{R}^n$. Como las filas de A se identifican con las columnas de A^T, podría escribirse también Col A^T en lugar de Fil A.

EJEMPLO 1 Sea

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix} \quad \text{y} \quad \begin{aligned} \mathbf{r}_1 &= (-2, -5, 8, 0, -17) \\ \mathbf{r}_2 &= (1, 3, -5, 1, 5) \\ \mathbf{r}_3 &= (3, 11, -19, 7, 1) \\ \mathbf{r}_4 &= (1, 7, -13, 5, -3) \end{aligned}$$

El espacio fila de A es el subespacio de $\mathbb{R}^5$ generado por $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. Esto es, Fil $A = \text{Gen}\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. Es natural que los vectores fila se escriban en forma horizontal, pero se podrían escribir como vectores columna si resultara más conveniente. ■

Si se conocieran algunas relaciones de dependencia lineal entre las filas de A dadas en el ejemplo 1, podría usarse el teorema del conjunto generador para reducir el conjunto generador a una base. Desafortunadamente, las operaciones por fila con A no proporcionan esa información, porque dichas operaciones alteran las relaciones de dependencia entre filas. Pero la reducción de A por filas siempre es valiosa, como lo muestra el teorema siguiente.

TEOREMA 13 Si dos matrices A y B son equivalentes por filas, entonces sus espacios de fila son iguales. Si B está en forma escalonada, entonces las filas distintas de cero de B constituyen una base para el espacio fila de A y para el de B.

DEMOSTRACIÓN Si B se obtiene a partir de A mediante operaciones por fila, las filas de B son combinaciones lineales de las filas de A. De modo que cualquier combinación lineal de las filas de B es automáticamente una combinación lineal de las filas de A. Entonces el espacio fila de B está incluido en el espacio fila de A. Como las operaciones por fila son reversibles, el mismo argumento muestra que el espacio fila de A es un subconjunto del espacio fila de B. Así que los dos espacios de filas son iguales. Si B está en forma escalonada, sus filas distintas de cero son linealmente independientes porque ninguna fila distinta de cero es una combinación lineal de las filas distintas de cero que están debajo. (Aplique el teorema 4 a las filas distintas de cero de B en orden inverso, con la primera fila al final.) Por lo tanto, las filas distintas de cero de B forman una base del espacio fila (común) de B y A. ■

El resultado principal de esta sección está relacionado con tres espacios: Fil A, Col A, y Nul A. El ejemplo siguiente prepara el camino para este resultado y muestra cómo *una* sucesión de operaciones por fila en A produce bases para los tres espacios.

EJEMPLO 2 Encuentre bases para el espacio fila, el espacio columna, y el espacio nulo de la matriz

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

Solución Para encontrar bases para el espacio fila y de columnas, reduzca por filas A a una forma escalonada:

$$A \sim B = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Por el teorema 13, las primeras tres filas de B forman una base para el espacio fila de A (y para el espacio fila de B). Entonces

$$\text{Base para Fil } A\text{: } \{(1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20)\}$$

Para el espacio columna, observe que en B los pivotes están en las columnas 1, 2 y 4. Por lo tanto, las columnas 1, 2 y 4 de A (no B) forman una base para Col A:

$$\text{Base para Col } A\text{: } \left\{ \begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix} \right\}$$

Observe que cualquier forma escalonada de A proporciona (con sus filas distintas de cero) una base para Fil A y también identifica las columnas pivote de A para Col A. Sin embargo, para Nul A, se necesita la *forma escalonada reducida.* Al realizar otras operaciones por fila sobre B, se obtiene

$$A \sim B \sim C = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

La ecuación $A\mathbf{x} = \mathbf{0}$ es equivalente a $C\mathbf{x} = \mathbf{0}$, esto es,

$$\begin{aligned} x_1 + x_3 + x_5 &= 0 \\ x_2 - 2x_3 + 3x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

Así $x_1 = -x_3 - x_5$, $x_2 = 2x_3 - 3x_5$, $x_4 = 5x_5$, con x_3 y x_5 como variables libres. Los cálculos usuales (analizados en la sección 4.2) muestran que

$$\text{Base para Nul } A: \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

Observe que, a diferencia de la base para Col A, las bases para Fil A y Nul A no tienen ninguna relación simple con las entradas de la propia A.[1]

Advertencia: Aunque en el ejemplo 2 las tres primeras filas de B son linealmente independientes, es erróneo concluir que las tres primeras filas de A son linealmente independientes. (De hecho, la tercera fila de A es dos veces la primera fila más siete veces la segunda fila.) Las operaciones por fila *no* conservan las relaciones de dependencia lineal entre las *filas* de una matriz.

El teorema del rango

El teorema siguiente describe relaciones fundamentales entre las dimensiones de Col A, Fil A y Nul A.

DEFINICIÓN

El **rango** de A es la dimensión del espacio columna de A.

Dado que Fil A es lo mismo que Col A^T, la dimensión del espacio fila de A es el rango de A^T. La dimensión del espacio nulo se denomina también como **nulidad** de A, pero no se usará este término en el texto.

Tal vez un lector atento ya haya descubierto cuando menos una parte del teorema siguiente, al realizar los ejercicios de la sección 4.5 o al leer el ejemplo 2 anterior.

TEOREMA 14

El teorema del rango

Las dimensiones del espacio columna y del espacio fila para una matriz A de $m \times n$ son iguales. Esta dimensión común, el rango de A, también es igual al número de posiciones pivote incluidas en A y satisface la ecuación

$$\text{rango } A + \dim \text{Nul } A = n$$

DEMOSTRACIÓN De acuerdo con el teorema 6 de la sección 4.3, rango A es el número de columnas pivote de A. De manera equivalente, rango A es el número de posiciones pivote en una forma escalonada B de A.

[1] Se puede encontrar una base para el espacio fila Fil A que use filas de A. Primero forme A^T, luego reduzca por filas hasta encontrar las columnas pivote de A^T. Estas columnas pivote de A^T son filas de A, y forman una base para el espacio fila de A.

Además, como B tiene una fila distinta de cero para cada pivote, y puesto que estas filas forman una base para el espacio fila de A, el rango de A también es la dimensión del espacio fila.

De acuerdo con la sección 4.5, la dimensión de Nul A es igual al número de variables libres incluidas en la ecuación $A\mathbf{x} = \mathbf{0}$. Dicho de otra manera, la dimensión de Nul A es el número de columnas de A que *no* son columnas pivote. (Es el número de estas columnas, no las propias columnas, lo que está relacionado con Nul A.) Desde luego,

$$\begin{Bmatrix} \text{número de} \\ \text{columnas pivote} \end{Bmatrix} + \begin{Bmatrix} \text{número de} \\ \text{columnas no pivote} \end{Bmatrix} = \begin{Bmatrix} \text{número de} \\ \text{columnas} \end{Bmatrix}$$

Esto demuestra el teorema. ■

Las ideas en que se basa el teorema 14 son perceptibles en los cálculos del ejemplo 2. Las tres posiciones pivote de la forma escalonada B determinan las variables básicas e identifican los tres vectores de la base para Col A y Fil A.

EJEMPLO 3

a. Si A es una matriz de 7×9 con un espacio nulo bidimensional, ¿cuál es el rango de A?

b. ¿Una matriz de 6×9 podría tener un espacio nulo bidimensional?

Solución

a. Dado que A tiene 9 columnas, $(\text{rango } A) + 2 = 9$ y, por lo tanto, rango $A = 7$.

b. No. Si una matriz de 6×9, llámese B, tuviera un espacio nulo bidimensional, tendría rango 7, por el teorema del rango. Pero las columnas de B son vectores en $\mathbb{R}^6$, así que la dimensión de Col B no puede ser mayor a 6; esto es, rango B no puede exceder de 6. ■

El siguiente ejemplo proporciona una buena manera de visualizar los subespacios que se han estado estudiando. En el capítulo 6 se aprenderá que Fil A y Nul A tienen sólo el vector cero en común, y que realmente son “perpendiculares” entre sí. El mismo hecho se aplica a Fil A^T (= Col A) y Nul A^T. Así que la figura del ejemplo 4 produce una buena imagen mental para el caso general. (La importancia de estudiar A^T junto con A se demuestra en el ejercicio 29.)

EJEMPLO 4 Sea $A = \begin{bmatrix} 3 & 0 & -1 \\ 3 & 0 & -1 \\ 4 & 0 & 5 \end{bmatrix}$. Se comprueba fácilmente que Nul A es el eje x_2, Fil A es el plano x_1x_3, Col A es el plano cuya ecuación es $x_1 - x_2 = 0$, y Nul A^T es el conjunto de todos los múltiplos de $(1, -1, 0)$. En la figura 1 se presentan Nul A y Fil A en el dominio de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$; el rango de esta función, Col A, se muestra en una copia aparte de $\mathbb{R}^3$, junto con Nul A^T. ■

Aplicaciones a sistemas de ecuaciones

El teorema del rango es una potente herramienta que sirve para procesar información acerca de sistemas de ecuaciones lineales. El siguiente ejemplo simula cómo se podría

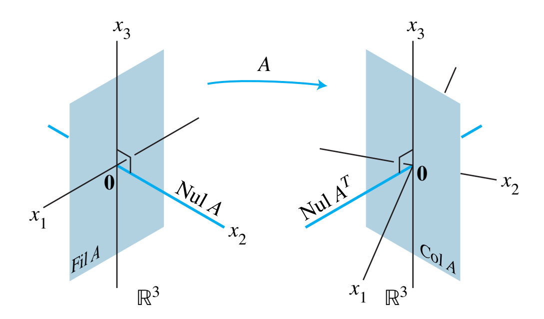


FIGURA 1 Subespacios asociados a una matriz A.

plantear un problema de la vida real por medio de ecuaciones lineales, sin mencionar explícitamente términos del álgebra lineal como matriz, subespacio o dimensión.

EJEMPLO 5 Un científico ha encontrado dos soluciones para un sistema homogéneo de 40 ecuaciones con 42 variables. Las dos soluciones no son múltiplos, y todas las demás soluciones pueden estructurarse al sumar múltiplos adecuados de estas dos soluciones. ¿Puede el científico estar *seguro* de que un sistema no homogéneo asociado (con los mismos coeficientes) tiene una solución?

Solución Sí. Sea A la matriz de coeficientes de 40×42 del sistema. La información proporcionada implica que las dos soluciones son linealmente independientes y generan Nul A. Así que dim Nul $A = 2$. Según el teorema del rango, dim Col $A = 42 - 2 = 40$. Como $\mathbb{R}^{40}$ es el único subespacio de $\mathbb{R}^{40}$ cuya dimensión es 40, Col A tiene que ser todo $\mathbb{R}^{40}$. Esto implica que toda ecuación no homogénea $A\mathbf{x} = \mathbf{b}$ tiene una solución.

Rango y el teorema de la matriz invertible

Los diversos conceptos de espacios vectoriales asociados a una matriz proporcionan varios enunciados adicionales al teorema de la matriz invertible. Aquí se incluyen solamente los nuevos enunciados, pero se hará referencia a ellos de modo que sigan a los enunciados del teorema de la matriz invertible original proporcionado en la sección 2.3.

TEOREMA

El teorema de la matriz invertible (continuación)

Sea A una matriz $n \times n$. Entonces los siguientes enunciados son equivalentes a la afirmación de que A es una matriz invertible.

m. Las columnas de A forman una base para $\mathbb{R}^n$.
n. Col $A = \mathbb{R}^n$.
o. dim Col $A = n$
p. rango $A = n$
q. Nul $A = \{\mathbf{0}\}$
r. dim Nul $A = 0$

DEMOSTRACIÓN El enunciado (m) es lógicamente equivalente a los enunciados (e) y (h) en lo que se refiere a la independencia lineal y la generación. Los otros cinco enunciados se vinculan al teorema por medio de la siguiente sucesión de implicaciones casi triviales:

$$(g) \Rightarrow (n) \Rightarrow (o) \Rightarrow (p) \Rightarrow (r) \Rightarrow (q) \Rightarrow (d)$$

El enunciado (g), el cual establece que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para cada $\mathbf{b}$ en $\mathbb{R}^n$, implica (n), porque Col A es precisamente el conjunto de todas las $\mathbf{b}$ tales que la ecuación $A\mathbf{x} = \mathbf{b}$ resulta ser consistente. Las implicaciones (n) $\Rightarrow$ (o) $\Rightarrow$ (p) provienen de las definiciones de dimensión y rango. Si el rango de A es n, el número de columnas de A, entonces dim Nul $A = 0$, de acuerdo con el teorema del rango, y así Nul $A = \{\mathbf{0}\}$. Por lo tanto (p) $\Rightarrow$ (r) $\Rightarrow$ (q). Además, (q) implica que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial, que es el enunciado (d). Como ya se sabe que los enunciados (d) y (g) son equivalentes al enunciado de que A es invertible, la demostración está completa. ■

SG Tabla ampliada para el TMI 4 a 21 (Expanded Table for the IMT 4-21)

Se ha evitado agregar al teorema de la matriz invertible enunciados evidentes acerca del espacio fila de A, porque este espacio es el espacio columna de A^T. Recuerde que, según (1) del teorema de la matriz invertible, A es invertible si, y sólo si, A^T es invertible. Por lo tanto, toda afirmación del teorema de la matriz invertible también puede emitirse acerca de A^T. ¡Esto duplicaría la longitud del teorema y produciría una lista de más de 30 enunciados!

NOTA NUMÉRICA

Muchos de los algoritmos que se analizan en este texto son útiles para entender conceptos y realizar cálculos sencillos a mano. Sin embargo, es muy común que estos algoritmos resulten inadecuados en problemas reales de gran envergadura.

La determinación del rango es un buen ejemplo. Parecería fácil reducir una matriz a su forma escalonada y contar los pivotes. Pero, a menos que se realice aritmética exacta con una matriz cuyas entradas se especifiquen con exactitud, las operaciones por fila pueden cambiar el rango aparente de una matriz. Por ejemplo, si el valor de x en la matriz $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ no se almacena con exactitud como 7 en una computadora, el rango puede ser 1 o 2, dependiendo de si la computadora trata a $x - 7$ como si fuera cero o no.

En aplicaciones prácticas, el rango efectivo de una matriz A se determina frecuentemente a partir de la descomposición en valores singulares de A, lo cual se estudiará en la sección 7.4. Esta descomposición también es una fuente confiable de bases para Col A, Fil A, Nul A, y Nul A^T.

CD El comando `rango` (The `rank` command)

PROBLEMAS DE PRÁCTICA

Las siguientes matrices son equivalentes por filas.

$$A = \begin{bmatrix} 2 & -1 & 1 & -6 & 8 \\ 1 & -2 & -4 & 3 & -2 \\ -7 & 8 & 10 & 3 & -10 \\ 4 & -5 & -7 & 0 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -2 & -4 & 3 & -2 \\ 0 & 3 & 9 & -12 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. Encuentre rango A y dim Nul A.
2. Encuentre bases para Col A y Fil A.
3. ¿Cuál sería el siguiente paso a realizar si se quisiera encontrar una base para Nul A?
4. ¿Cuántas columnas pivote hay en una forma escalonada de A^T?

4.6 EJERCICIOS

En los ejercicios 1 a 4, suponga que la matriz A es equivalente por filas a B. Sin realizar cálculos, enliste rango A y dim Nul A. Después encuentre bases para Col A, Fil A, y Nul A.

1. $A = \begin{bmatrix} 1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & -3 & 4 & -1 & 9 \\ -2 & 6 & -6 & -1 & -10 \\ -3 & 9 & -6 & -6 & -3 \\ 3 & -9 & 4 & 9 & 0 \end{bmatrix},$

$B = \begin{bmatrix} 1 & -3 & 0 & 5 & -7 \\ 0 & 0 & 2 & -3 & 8 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ -2 & 3 & -3 & -3 & -4 \\ 4 & -6 & 9 & 5 & 9 \\ -2 & 3 & 3 & -4 & 1 \end{bmatrix},$

$B = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ 0 & 0 & 3 & -1 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 1 & 2 & -4 & 10 & 13 & -12 \\ 1 & -1 & -1 & 1 & 1 & -3 \\ 1 & -3 & 1 & -5 & -7 & 3 \\ 1 & -2 & 0 & 0 & -5 & -4 \end{bmatrix},$

$B = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 0 & 1 & -1 & 3 & 4 & -3 \\ 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

5. Si una matriz A de 3×8 tiene rango 3, encuentre dim Nul A, dim Fil A, y rango A^T.

6. Si una matriz A de 6×3 tiene rango 3, encuentre dim Nul A, dim Fil A, y rango A^T.

7. Suponga que una matriz A de 4×7 tiene cuatro columnas pivote. ¿Es Col $A = \mathbb{R}^4$? ¿Es Nul $A = \mathbb{R}^3$? Explique sus respuestas.

8. Suponga que una matriz A de 5×6 tiene cuatro columnas pivote. ¿Cuál es el valor de dim Nul A? ¿Es Col $A = \mathbb{R}^4$? ¿Por qué sí o por qué no?

9. Si el espacio nulo de una matriz A de 5×6 es de dimensión 4, ¿cuál es la dimensión del espacio columna de A?

10. Si el espacio nulo de una matriz A de 7×6 es de dimensión 5, ¿cuál es la dimensión del espacio columna de A?

11. Si el espacio nulo de una matriz A de 8×5 es de dimensión 2, ¿cuál es la dimensión del espacio fila de A?

12. Si el espacio nulo de una matriz A de 5×6 es de dimensión 4, ¿cuál es la dimensión del espacio fila de A?

13. Si A es una matriz de 7×5, ¿cuál es el mayor valor posible para el rango de A? Si A es una matriz de 5×7, ¿cuál es el máximo valor posible para el rango de A? Explique sus respuestas.

14. Si A es una matriz de 4×3, ¿cuál es la mayor dimensión posible para el espacio fila de A? Si A es una matriz de 3×4, ¿cuál es la mayor dimensión posible para el espacio fila de A? Explique sus respuestas.

15. Si A es una matriz de 6×8, ¿cuál es la menor dimensión posible para Nul A?

16. Si A es una matriz de 6×4, ¿cuál es la menor dimensión posible para Nul A?

En los ejercicios 17 y 18, A es una matriz de $m \times n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

17. a. El espacio fila de A es lo mismo que el espacio columna de A^T.

b. Si B es cualquier forma escalonada de A y tiene tres filas distintas de cero, entonces las primeras tres filas de A forman una base para Fil A.

c. Las dimensiones del espacio fila y del espacio columna de A son las mismas, aunque A no sea cuadrada.

d. La suma de las dimensiones del espacio fila y del espacio nulo de A es igual al número de filas incluidas en A.

e. En una computadora, las operaciones por fila pueden alterar el rango aparente de una matriz.

18. a. Si B es cualquier forma escalonada de A, entonces las columnas pivote de B forman una base para el espacio columna de A.

b. Las operaciones por fila preservan las relaciones de dependencia lineal entre las filas de A.

c. La dimensión del espacio nulo de A es el número de columnas de A que *no* son columnas pivote.

d. El espacio fila de A^T es lo mismo que el espacio columna de A.

e. Si A y B son equivalentes por filas, entonces sus espacios de filas son iguales.

19. Suponga que las soluciones de un sistema homogéneo de cinco ecuaciones lineales con seis incógnitas son todas múltiplos de una solución distinta de cero. ¿El sistema tendrá, necesariamente, una solución para todas las posibles selecciones de constantes en los miembros derechos de las ecuaciones? Explique su respuesta.

20. Suponga que un sistema no homogéneo de seis ecuaciones con ocho incógnitas tiene una solución, con dos variables libres. ¿Pueden cambiarse algunas constantes en los miembros derechos de las ecuaciones de tal manera que el nuevo sistema resulte inconsistente? Explique su respuesta.

21. Suponga que un sistema no homogéneo de nueve ecuaciones lineales con diez incógnitas tiene una solución para todas las posibles constantes de los miembros derechos de las ecuaciones. ¿Pueden encontrarse dos soluciones distintas de cero del sistema homogéneo asociado que *no* sean múltiplos una de la otra? Analice el planteamiento.

22. ¿Es posible que todas las soluciones de un sistema homogéneo de diez ecuaciones lineales con doce variables sean múltiplos de una solución fija distinta de cero? Analice el planteamiento.

23. Un sistema homogéneo de doce ecuaciones lineales con ocho incógnitas tiene dos soluciones fijas que no son múltiplos una de la otra y cualquier otra solución es una combinación lineal de estas dos soluciones. ¿Es posible describir al conjunto de todas las soluciones con menos de doce ecuaciones homogéneas? Si la respuesta es sí, ¿con cuántas ecuaciones? Analice el planteamiento.

24. ¿Es posible que un sistema no homogéneo de siete ecuaciones con seis incógnitas tenga una solución única para algún conjunto de constantes del miembro derecho? Es posible que un sistema como éste pueda tener una solución única para cada miembro derecho? Explique sus respuestas.

25. Un científico resuelve un sistema no homogéneo de diez ecuaciones lineales con doce incógnitas y encuentra que tres de las incógnitas son variables libres. ¿Puede el científico estar seguro de que, si los miembros derechos de las ecuaciones se cambian, el nuevo sistema no homogéneo tendrá una solución? Analice el planteamiento.

26. En teoría estadística, un requisito frecuente es que una matriz sea de *rango pleno.* Esto es, el rango debe ser lo más grande posible. Explique por qué una matriz $m \times n$ con más filas que columnas tiene rango pleno si, y sólo si, sus columnas son linealmente independientes.

Los ejercicios 27 a 29 se refieren a una matriz A $m \times n$ y a lo que suele denominarse *subespacios fundamentales* determinados por A.

27. ¿Cuáles de los siguientes subespacios, Fil A, Col A, Nul A, Fil A^T, Col A^T, y Nul A^T están en $\mathbb{R}^m$ y cuáles están en $\mathbb{R}^n$ ¿Cuántos subespacios distintos hay en esta lista?

28. Justifique las siguientes igualdades:

a. $\dim \text{Fil } A + \dim \text{Nul } A = n$ Número de columnas de A

b. $\dim \text{Col } A + \dim \text{Nul } A^T = m$ Número de filas de A

29. Use el ejercicio 28 para explicar por qué la ecuación tiene una solución para toda $\mathbf{b}$ en $\mathbb{R}^m$ si, y sólo si, la ecuación $A^T\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial.

30. Suponga que A es una matriz de $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$. ¿Qué debe ser cierto acerca de los dos números rango $[A \quad \mathbf{b}]$ y rango A para que la ecuación $A\mathbf{x} = \mathbf{b}$ sea consistente?

Las matrices con rango 1 son importantes para muchos algoritmos de computadora y en varios contextos teóricos, incluyendo la descomposición en valores singulares del capítulo 7. Puede demostrarse que una matriz A de $m \times n$ tiene rango 1 si, y sólo si, es un producto exterior; es decir, $A = \mathbf{u}\mathbf{v}^T$ para algunas $\mathbf{u}$ en $\mathbb{R}^m$ y $\mathbf{v}$ en $\mathbb{R}^n$. Los ejercicios 31 a 33 sugieren por qué esta propiedad es cierta.

31. Compruebe que rango $\mathbf{u}\mathbf{v}^T \leq 1$ si $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

32. Sea $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Encuentre $\mathbf{v}$ en $\mathbb{R}^3$ tal que $\begin{bmatrix} 1 & -3 & 4 \\ 2 & -6 & 8 \end{bmatrix} = \mathbf{u}\mathbf{v}^T$.

33. Sea A cualquier matriz de 2×3 tal que rango $A = 1$, sea $\mathbf{u}$ la primera columna de A, y suponga que $\mathbf{u} \neq \mathbf{0}$. Explique por qué hay un vector $\mathbf{v}$ en $\mathbb{R}^3$ tal que $A = \mathbf{u}\mathbf{v}^T$. ¿Cómo podría modificarse esta construcción si la primera columna de A fuera cero?

34. Sean A una matriz $m \times n$ de rango $r > 0$ y U una forma escalonada de A. Explique por qué existe una matriz invertible E tal que $A = EU$, y use esta factorización para escribir A como la suma de r matrices con rango 1. [*Sugerencia:* Vea el teorema 10 de la sección 2.4.]

35. **[M]** Sea $A = \begin{bmatrix} 7 & -9 & -4 & 5 & 3 & -3 & -7 \\ -4 & 6 & 7 & -2 & -6 & -5 & 5 \\ 5 & -7 & -6 & 5 & -6 & 2 & 8 \\ -3 & 5 & 8 & -1 & -7 & -4 & 8 \\ 6 & -8 & -5 & 4 & 4 & 9 & 3 \end{bmatrix}$.

a. Construya matrices C y N cuyas columnas sean bases para Col A y Nul A, respectivamente, y construya una matriz R cuyas filas formen una base para Fil A.

b. Construya una matriz M cuyas columnas formen una base para Nul A^T, forme las matrices $S = [R^T \quad N]$ y $T = [C \quad M]$, y explique por qué S y T deben ser cuadradas. Compruebe que tanto S como T son invertibles.

36. **[M]** Repita el ejercicio 35 para una matriz aleatoria A de 6×7 con valores enteros y rango de 4 o menos. Una manera de construir A es creando una matriz aleatoria J de 6×4 y una matriz aleatoria con valores enteros K de 4×7, y estableciendo $A = JK$. [Vea el ejercicio suplementario 12 que aparece al final del capítulo; y vea la *guía de estudio* (*Study Guide*) para programas de generación de matrices.]

37. **[M]** Sea A la matriz del ejercicio 35. Construya una matriz C cuyas columnas sean las columnas pivote de A, y estructure una matriz R cuyas filas sean las filas distintas de cero de la forma escalonada reducida de A. Calcule CR y analice lo que encuentre.

38. **[M]** Repita el ejercicio 37 para tres matrices aleatorias A de 5×7 con valores enteros cuyos rangos sean 5, 4 y 3. Formule un supuesto acerca de cómo está relacionada CR con A para cualquier matriz A. Demuestre su conjetura.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. A tiene dos columnas pivote, así que rango $A = 2$. Como A tiene cinco columnas en total, dim Nul $A = 5 - 2 = 3$.

2. Las columnas pivote de A son las primeras dos columnas. Por lo tanto, una base para Col A es

$$\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 8 \\ -5 \end{bmatrix} \right\}$$

En B, las filas distintas de cero forman una base para Fil A, a saber, $\{(1, -2, -4, 3, -2), (0, 3, 9, -12, 12)\}$. En este ejemplo en particular, sucede que cualesquiera dos filas de A forman una base para el espacio fila, porque el espacio fila es de dimensión dos y ninguna de las filas de A es múltiplo de otra fila. En general, deben usarse las filas diferentes de cero en una forma escalonada de A como base para Fil A, no las filas de la propia A.

SG Revisión principal de los conceptos clave, 4 a 24 (Major Review of Key Concepts 4-24)

3. Para Nul A, el siguiente paso es realizar operaciones por fila con B para obtener la forma escalonada reducida de A.

4. Rango A^T = rango A, de acuerdo con el teorema del rango, porque Col A^T = Fil A. Así, A^T tiene dos posiciones pivote.

4.7 CAMBIO DE BASE

Cuando se elige una base $\mathcal{B}$ para un espacio vectorial V de dimensión n, la función de coordenadas asociada sobre $\mathbb{R}^n$ proporciona un sistema de coordenadas para V. Cada $\mathbf{x}$ en V se identifica de manera única con su vector de $\mathcal{B}$-coordenadas $[\mathbf{x}]_{\mathcal{B}}$.[1]

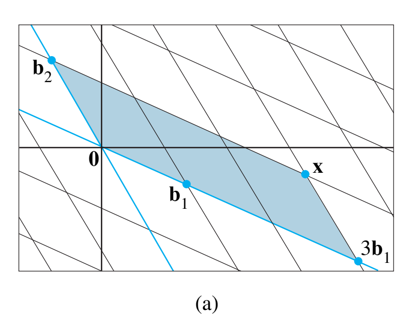
En algunas aplicaciones, inicialmente se describe un problema usando una base $\mathcal{B}$, pero la solución del problema se facilita al cambiar la base $\mathcal{B}$ a una nueva base $\mathcal{C}$. (En los capítulos 5 y 7 se darán ejemplos de esto.) A cada vector se le asigna un nuevo vector de $\mathcal{C}$-coordenadas. En esta sección se estudiará cómo están relacionados $[\mathbf{x}]_{\mathcal{C}}$ y $[\mathbf{x}]_{\mathcal{B}}$ para cada $\mathbf{x}$ en V.

[1]Piense en $[\mathbf{x}]_{\mathcal{B}}$ como un "nombre" para $\mathbf{x}$ que enlista los pesos usados al construir $\mathbf{x}$ como una combinación lineal de los vectores de base en $\mathcal{B}$.

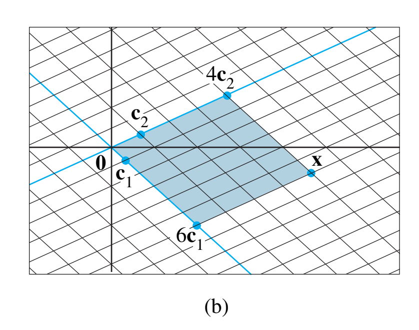
Para visualizar el problema, considere el sistema de dos coordenadas mostrado en la figura 1. En la figura l(a), $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$, mientras que en la figura 1(b), la misma $\mathbf{x}$ se muestra como $\mathbf{x} = 6\mathbf{c}_1 + 4\mathbf{c}_2$. Esto es,

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{y} \quad [\,\mathbf{x}\,]_{\mathcal{C}} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

El problema consiste en encontrar la relación que hay entre los dos vectores de coordenadas. En el ejemplo 1 se muestra cómo hacer esto, suponiendo que ya se sabe cómo están formadas $\mathbf{b}_1$ y $\mathbf{b}_2$ a partir de $\mathbf{c}_1$ y $\mathbf{c}_2$.



(a)



(b)

FIGURA 1 Dos sistemas de coordenadas para el mismo espacio vectorial.

EJEMPLO 1 Considere dos bases $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ para un espacio vectorial V, tales que

$$\mathbf{b}_1 = 4\mathbf{c}_1 + \mathbf{c}_2 \quad \text{y} \quad \mathbf{b}_2 = -6\mathbf{c}_1 + \mathbf{c}_2 \tag{1}$$

Suponga que

$$\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2 \tag{2}$$

Esto es, suponga que $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Encuentre $[\,\mathbf{x}\,]_{\mathcal{C}}$.

Solución Aplique la función de coordenadas determinada mediante $\mathcal{C}$ a $\mathbf{x}$ en (2). Como la función de coordenadas es una transformación lineal,

$$\begin{aligned} [\,\mathbf{x}\,]_{\mathcal{C}} &= [\,3\mathbf{b}_1 + \mathbf{b}_2\,]_{\mathcal{C}} \\ &= 3[\,\mathbf{b}_1\,]_{\mathcal{C}} + [\,\mathbf{b}_2\,]_{\mathcal{C}} \end{aligned}$$

Esta ecuación vectorial puede escribirse como una ecuación matricial, usando los vectores de la combinación lineal como las columnas de una matriz:

$$[\,\mathbf{x}\,]_{\mathcal{C}} = \begin{bmatrix} [\,\mathbf{b}_1\,]_{\mathcal{C}} & [\,\mathbf{b}_2\,]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{3}$$

De esta fórmula se obtiene $[\mathbf{x}]_{\mathcal{C}}$, una vez que se conocen las columnas de la matriz. De (1),

$$[\,\mathbf{b}_1\,]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad \text{y} \quad [\,\mathbf{b}_2\,]_{\mathcal{C}} = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$$

Entonces (3) proporciona la solución:

$$[\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Las $\mathcal{C}$-coordenadas de $\mathbf{x}$ coinciden con las de $\mathbf{x}$ en la figura 1.

El argumento usado para deducir la fórmula (3) se generaliza para producir el siguiente resultado. (Vea los ejercicios 15 y 16.)

TEOREMA 15

Sean $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ las bases de un espacio vectorial V. Entonces existe una sola matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ de $n \times n$ tal que

$$[\mathbf{x}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}} \tag{4}$$

Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son los vectores de $\mathcal{C}$-coordenadas de los vectores de la base $\mathcal{B}$. Esto es,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} & \cdots & [\mathbf{b}_n]_{\mathcal{C}} \end{bmatrix} \tag{5}$$

La matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ del teorema 15 se denomina **matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$**. La multiplicación por $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ convierte $\mathcal{B}$-coordenadas en $\mathcal{C}$-coordenadas.[2] En la figura 2 se ilustra la ecuación de cambio de coordenadas (4).

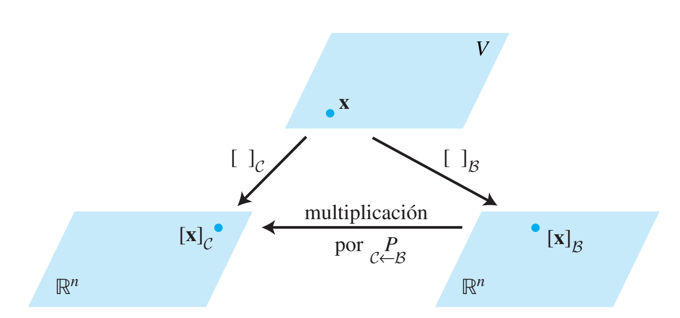


FIGURA 2 Dos sistemas de coordenadas para V.

Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son linealmente independientes porque son los vectores de coordenadas del conjunto linealmente independiente $\mathcal{B}$. (Vea el ejercicio 25 de la sección 4.4.) Como $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es cuadrada, debe ser invertible, según el teorema de la matriz invertible. Al multiplicar por la izquierda ambos lados de (4) por $(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$ se obtiene

$$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}[\mathbf{x}]_{\mathcal{C}} = [\mathbf{x}]_{\mathcal{B}}$$

[2]Para recordar la forma de construir la matriz, piense en $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}}$ como una combinación lineal de las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. El producto matriz-vector es un vector de $\mathcal{C}$-coordenadas, así que las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ deben ser también vectores de $\mathcal{C}$-coordenadas.

Entonces $(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$ es la matriz que convierte $\mathcal{C}$-coordenadas en $\mathcal{B}$-coordenadas. Esto es,

$$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P} \tag{6}$$

Cambio de base en $\mathbb{R}^n$

Si $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $\mathcal{E}$ es la *base estándar* $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ en $\mathbb{R}^n$, entonces $[\mathbf{b}_1]_{\mathcal{E}} = \mathbf{b}_1$, y lo mismo es válido para los otros vectores en $\mathcal{B}$. En este caso, $\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}$ es igual a la matriz de cambio de coordenadas $P_{\mathcal{B}}$ que se introdujo en la sección 4.4, a saber,

$$P_{\mathcal{B}} = [\, \mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n \,]$$

Para cambiar coordenadas entre dos bases no estándar en $\mathbb{R}^n$, se requiere el teorema 15. El teorema establece que, para resolver el problema de cambio de coordenadas, se necesitan los vectores de coordenadas de la base anterior relativos a la base nueva.

EJEMPLO 2 Sean $\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, y considere las bases para $\mathbb{R}^2$ dadas por $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$. Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

Solución La matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ incluye los vectores de $\mathcal{C}$-coordenadas de $\mathbf{b}_1$ y $\mathbf{b}_2$. Sea $[\mathbf{b}_1]_{\mathcal{C}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ y $[\mathbf{b}_2]_{\mathcal{C}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Entonces, por definición,

$$[\, \mathbf{c}_1 \quad \mathbf{c}_2 \,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}_1 \quad \text{y} \quad [\, \mathbf{c}_1 \quad \mathbf{c}_2 \,] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{b}_2$$

Para resolver simultáneamente ambos sistemas, aumente la matriz de coeficientes con $\mathbf{b}_1$ *y* con $\mathbf{b}_2$ y reduzca por filas:

$$[\, \mathbf{c}_1 \quad \mathbf{c}_2 \mid \mathbf{b}_1 \quad \mathbf{b}_2 \,] = \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right] \tag{7}$$

Entonces,

$$[\, \mathbf{b}_1 \,]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \quad \text{y} \quad [\, \mathbf{b}_2 \,]_{\mathcal{C}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

La matriz de cambio de coordenadas deseada es, por lo tanto,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \big[\, [\, \mathbf{b}_1 \,]_{\mathcal{C}} \quad [\, \mathbf{b}_2 \,]_{\mathcal{C}} \,\big] = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$$

Observe que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ del ejemplo 2 aparece ya en (7). Esto no es sorprendente porque la primera columna de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es el resultado de reducir por filas $[\, \mathbf{c}_1 \quad \mathbf{c}_2 \mid \mathbf{b}_1 \,]$ a $[\, I \mid [\, \mathbf{b}_1 \,]_{\mathcal{C}} \,]$, y lo mismo es válido para la segunda columna de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. Entonces,

$$[\, \mathbf{c}_1 \quad \mathbf{c}_2 \mid \mathbf{b}_1 \quad \mathbf{b}_2 \,] \sim [\, I \mid \underset{\mathcal{C}\leftarrow\mathcal{B}}{P} \,]$$

Un procedimiento análogo sirve para encontrar la matriz de cambio de coordenadas entre dos bases cualesquiera en $\mathbb{R}^n$.

EJEMPLO 3 Sean $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$, y considere las bases para $\mathbb{R}^2$ dadas por $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$.

a. Encuentre la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

b. Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

Solución

a. Observe que se necesita $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ en lugar de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, y calcule

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \vdots & \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} = \left[\begin{array}{rr:rr} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array}\right] \sim \left[\begin{array}{rr:rr} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array}\right]$$

Así,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}$$

b. De acuerdo con el inciso (a) y la propiedad (6) anterior, (intercambiando $\mathcal{B}$ y $\mathcal{C}$).

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = (\underset{\mathcal{B}\leftarrow\mathcal{C}}{P})^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix}$$

Otra descripción de la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ usa las matrices de cambio de coordenadas $P_\mathcal{B}$ y $P_\mathcal{C}$ que convierten $\mathcal{B}$-coordenadas y $\mathcal{C}$-coordenadas, respectivamente, en coordenadas estándar. Recuerde que para cada $\mathbf{x}$ en $\mathbb{R}^n$,

$$P_\mathcal{B}[\mathbf{x}]_\mathcal{B} = \mathbf{x}, \quad P_\mathcal{C}[\mathbf{x}]_\mathcal{C} = \mathbf{x}, \quad \text{y} \quad [\mathbf{x}]_\mathcal{C} = P_\mathcal{C}^{-1}\mathbf{x}$$

Entonces,

$$[\mathbf{x}]_\mathcal{C} = P_\mathcal{C}^{-1}\mathbf{x} = P_\mathcal{C}^{-1}P_\mathcal{B}[\mathbf{x}]_\mathcal{B}$$

En $\mathbb{R}^n$, la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ puede calcularse como $P_\mathcal{C}^{-1}P_\mathcal{B}$. En realidad, para matrices más grandes que 2×2, un algoritmo análogo al del ejemplo 3 es más rápido que calcular $P_\mathcal{C}^{-1}$ y después $P_\mathcal{C}^{-1}P_\mathcal{B}$. Vea el ejercicio 12 en la sección 2.2.

PROBLEMAS DE PRÁCTICA

1. Sean $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2\}$ y $\mathcal{G} = \{\mathbf{g}_1, \mathbf{g}_2\}$ bases para un espacio vectorial V, y sea P una matriz cuyas columnas son $[\mathbf{f}_1]_\mathcal{G}$ y $[\mathbf{f}_2]_\mathcal{G}$. ¿Cuál de las siguientes ecuaciones es satisfecha por P para toda $\mathbf{v}$ en V?

$$\text{(i)}\ [\,\mathbf{v}\,]_\mathcal{F} = P[\,\mathbf{v}\,]_\mathcal{G} \qquad \text{(ii)}\ [\,\mathbf{v}\,]_\mathcal{G} = P[\,\mathbf{v}\,]_\mathcal{F}$$

2. Sean $\mathcal{B}$ y $\mathcal{C}$ como en el ejemplo 1. Use los resultados de ese ejemplo para encontrar la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

4.7 EJERCICIOS

1. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para un espacio vectorial V, y suponga que $\mathbf{b}_1 = 6\mathbf{c}_1 - 2\mathbf{c}_2$ y $\mathbf{b}_2 = 9\mathbf{c}_1 - 4\mathbf{c}_2$.

a. Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

b. Encuentre $[\mathbf{x}]_{\mathcal{C}}$ para $\mathbf{x} = -3\mathbf{b}_1 + 2\mathbf{b}_2$. Use el inciso (a).

2. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para un espacio vectorial V, y suponga que $\mathbf{b}_1 = -\mathbf{c}_1 + 4\mathbf{c}_2$ y $\mathbf{b}_2 = 5\mathbf{c}_1 - 3\mathbf{c}_2$.

a. Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

b. Encuentre $[\mathbf{x}]_{\mathcal{C}}$ para $\mathbf{x} = 5\mathbf{b}_1 + 3\mathbf{b}_2$.

3. Sean $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2\}$ y $\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2\}$ bases para V, y sea P una matriz cuyas columnas son $[\mathbf{u}_1]_{\mathcal{W}}$ y $[\mathbf{u}_2]_{\mathcal{W}}$. ¿Cuál de las siguientes ecuaciones es satisfecha por P para toda $\mathbf{x}$ en V?

(i) $[\mathbf{x}]_{\mathcal{U}} = P[\mathbf{x}]_{\mathcal{W}}$ (ii) $[\mathbf{x}]_{\mathcal{W}} = P[\mathbf{x}]_{\mathcal{U}}$

4. Sean $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ bases para V, y sea $P = [[\mathbf{d}_1]_{\mathcal{A}} \quad [\mathbf{d}_2]_{\mathcal{A}} \quad [\mathbf{d}_3]_{\mathcal{A}}]$. ¿Cuál de las siguientes ecuaciones es satisfecha por P para toda $\mathbf{x}$ en V?

(i) $[\mathbf{x}]_{\mathcal{A}} = P[\mathbf{x}]_{\mathcal{D}}$ (ii) $[\mathbf{x}]_{\mathcal{D}} = P[\mathbf{x}]_{\mathcal{A}}$

5. Sean $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ bases para un espacio vectorial V, y suponga que $\mathbf{a}_1 = 4\mathbf{b}_1 - \mathbf{b}_2$, $\mathbf{a}_2 = -\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3$, y $\mathbf{a}_3 = \mathbf{b}_2 - 2\mathbf{b}_3$.

a. Encuentre la matriz de cambio de coordenadas de $\mathcal{A}$ a $\mathcal{B}$.

b. Encuentre $[\mathbf{x}]_{\mathcal{B}}$ para $\mathbf{x} = 3\mathbf{a}_1 + 4\mathbf{a}_2 + \mathbf{a}_3$.

6. Sean $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ y $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ bases para un espacio vectorial V, y suponga que $\mathbf{f}_1 = 2\mathbf{d}_1 - \mathbf{d}_2 + \mathbf{d}_3$, $\mathbf{f}_2 = 3\mathbf{d}_2 + \mathbf{d}_3$, y $\mathbf{f}_3 = -3\mathbf{d}_1 + 2\mathbf{d}_3$.

a. Encuentre la matriz de cambio de coordenadas de $\mathcal{F}$ a $\mathcal{D}$.

b. Encuentre $[\mathbf{x}]_{\mathcal{D}}$ para $\mathbf{x} = \mathbf{f}_1 - 2\mathbf{f}_2 + 2\mathbf{f}_3$.

En los ejercicios 7 a 10, sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para $\mathbb{R}^2$. En cada ejercicio, encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$ y la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

7. $\mathbf{b}_1 = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

8. $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 8 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

9. $\mathbf{b}_1 = \begin{bmatrix} -6 \\ -1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$

10. $\mathbf{b}_1 = \begin{bmatrix} 7 \\ -2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$

En los ejercicios 11 y 12, $\mathcal{B}$ y $\mathcal{C}$ son bases para un espacio vectorial V. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

11. a. Las columnas de la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son vectores de $\mathcal{B}$-coordenadas de los vectores en $\mathcal{C}$.

b. Si $V = \mathbb{R}^n$ y $\mathcal{C}$ es la base *estándar* para V, entonces $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es igual a la matriz de cambio de coordenadas $P_{\mathcal{B}}$ que se introdujo en la sección 4.4.

12. a. Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son linealmente independientes.

b. Si $V = \mathbb{R}^2$, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$, entonces la reducción por filas de $[\mathbf{c}_1 \quad \mathbf{c}_2 \quad \mathbf{b}_1 \quad \mathbf{b}_2]$ a $[I \quad P]$ produce una matriz P que satisface $[\mathbf{x}]_{\mathcal{B}} = P[\mathbf{x}]_{\mathcal{C}}$ para toda $\mathbf{x}$ en V.

13. En $\mathbb{P}_2$, encuentre la matriz de cambio de coordenadas de la base $\mathcal{B} = \{1 - 2t + t^2, 3 - 5t + 4t^2, 2t + 3t^2\}$ a la base estándar $\mathcal{C} = \{1, t, t^2\}$. Después encuentre el vector de $\mathcal{B}$-coordenadas para $-1 + 2t$.

14. En $\mathbb{P}_2$, encuentre la matriz de cambio de coordenadas de la base $\mathcal{B} = \{1 - 3t^2, 2 + t - 5t^2, 1 + 2t\}$ a la base estándar. Después escriba t^2 como una combinación lineal de los polinomios en $\mathcal{B}$.

En los ejercicios 15 y 16 proporcione una prueba del teorema 15. Complete la justificación para cada paso.

15. Dado $\mathbf{v}$ en V, existen escalares $x_1, \ldots, x_n$, tales que

$$\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n$$

porque (a)__________. Aplique la función de coordenadas determinada mediante la base $\mathcal{C}$, y obtenga

$$[\mathbf{v}]_{\mathcal{C}} = x_1[\mathbf{b}_1]_{\mathcal{C}} + x_2[\mathbf{b}_2]_{\mathcal{C}} + \cdots + x_n[\mathbf{b}_n]_{\mathcal{C}}$$

porque (b)__________. Esta ecuación puede escribirse en la forma

$$[\mathbf{v}]_{\mathcal{C}} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} & \cdots & [\mathbf{b}_n]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad (8)$$

por la definición de (c)____________. Esto muestra que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ presentada en (5) satisface $[\mathbf{v}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{v}]_{\mathcal{B}}$ para cada $\mathbf{v}$ en V, porque el vector en el miembro derecho de (8) es (d)___________.

16. Suponga que Q es cualquier matriz tal que

$$[\mathbf{v}]_{\mathcal{C}} = Q[\mathbf{v}]_{\mathcal{B}} \quad \text{para cada } \mathbf{v} \text{ en } V \qquad (9)$$

Establezca $\mathbf{v} = \mathbf{b}_1$ en (9). Entonces (9) muestra que $[\mathbf{b}_1]_{\mathcal{C}}$ es la primera columna de Q porque (a)__________. De manera similar, para $k = 2, \ldots, n$, la k-ésima columna de Q es (b)___________ porque (c)___________. Esto muestra que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ definida mediante (5) en el teorema 15 es la única que satisface la condición 4.

17. **[M]** Sea $\mathcal{B} = \{\mathbf{x}_0, \ldots, \mathbf{x}_6\}$ y $\mathcal{C} = \{\mathbf{y}_0, \ldots, \mathbf{y}_6\}$, donde $\mathbf{x}_k$ es la función $\cos^k t$ y $\mathbf{y}_k$ es la función $\cos kt$. En el ejercicio 34 de la sección 4.5 se mostró que tanto $\mathcal{B}$ como $\mathcal{C}$ son bases para el espacio vectorial $H = \text{Gen}\,\{\mathbf{x}_0, \ldots, \mathbf{x}_6\}$.

a. Establezca $P = \begin{bmatrix} [\mathbf{y}_0]_{\mathcal{B}} & \cdots & [\mathbf{y}_6]_{\mathcal{B}} \end{bmatrix}$, y calcule P^{-1}.

b. Explique por qué las columnas de P^{-1} son los vectores de $\mathcal{C}$-coordenadas de $\mathbf{x}_0, \ldots, \mathbf{x}_6$. Después use estos vectores de coordenadas para escribir identidades trigonométricas que expresen potencias de $\cos t$ en términos de las funciones en $\mathcal{C}$.

18. **[M]** (*Se requiere cálculo.*)[3] Recuerde del cálculo que las integrales como

$$\int (5\cos^3 t - 6\cos^4 t + 5\cos^5 t - 12\cos^6 t)\,dt \qquad (10)$$

son tediosas de calcular. (El método acostumbrado es aplicar repetidamente integración por partes y utilizar la fórmula para la mitad del ángulo.) Use las matrices P o P^{-1} del ejercicio 17 para transformar (10); luego determine la integral.

[3] La idea para los ejercicios 17 y 18, y para cinco ejercicios relacionados que aparecen en secciones anteriores, proviene de un documento de Jack W. Rogers, Jr., de Auburn University, el cual fue presentado en una reunión de la International Linear Algebra Society, en agosto de 1995. Vea "Applications of Linear Algebra in Calculus", *American Mathematical Monthly* **104** (1), 1997.

19. **[M]** Sean

$$P = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -5 & 0 \\ 4 & 6 & 1 \end{bmatrix},$$

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -8 \\ 5 \\ 2 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -7 \\ 2 \\ 6 \end{bmatrix}$$

a. Encuentre una base $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ para $\mathbb{R}^3$ tal que P sea la matriz de cambio de coordenadas de $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ a la base $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. [*Pista:* ¿Qué representan las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$?]

b. Encuentre una base $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ para $\mathbb{R}^3$ tal que P sea la matriz de cambio de coordenadas de $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a la base $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

20. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$, y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ bases para un espacio vectorial bidimensional.

a. Escriba una ecuación que relacione las matrices $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, $\underset{\mathcal{D}\leftarrow\mathcal{C}}{P}$, y $\underset{\mathcal{D}\leftarrow\mathcal{B}}{P}$. Justifique el resultado.

b. **[M]** Use un programa de matrices para encontrar la ecuación con mayor facilidad o para comprobar la que usted haya escrito. Trabaje con tres bases para $\mathbb{R}^2$. (Vea los ejercicios 7 a 10.)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Como las columnas de P son vectores de $\mathcal{G}$-coordenadas, un vector de la forma $P\mathbf{x}$ debe ser un vector de $\mathcal{G}$-coordenadas. Entonces P satisface la ecuación (ii).
2. Los vectores de coordenadas encontrados en el ejemplo 1 muestran que

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\,\mathbf{b}_1\,]_{\mathcal{C}} & [\,\mathbf{b}_2\,]_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}$$

Por lo tanto,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = (\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1} = \frac{1}{10}\begin{bmatrix} 1 & 6 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} .1 & .6 \\ -.1 & .4 \end{bmatrix}$$

4.8 APLICACIONES A ECUACIONES EN DIFERENCIAS

En la actualidad abundan las computadoras potentes, y cada vez más problemas científicos y de ingeniería se tratan utilizando sólo datos discretos, o digitales, en lugar de datos continuos. Las ecuaciones en diferencias son a menudo la herramienta adecuada para analizar tales datos. Aunque se use una ecuación diferencial para modelar un proceso continuo, frecuentemente se emplea una ecuación en diferencias relacionada para producir una solución numérica.

En esta sección se destacan algunas propiedades fundamentales de las ecuaciones en diferencias que se explican mejor mediante el uso de álgebra lineal.

Señales en tiempo discreto

El espacio vectorial $\mathbb{S}$ de las señales en tiempo discreto se introdujo en la sección 4.1. Una **señal** en $\mathbb{S}$ es una función definida sólo en los enteros y puede visualizarse como una sucesión de números, por ejemplo, $\{y_k\}$. En la figura 1 se muestran tres señales típicas cuyos términos generales son $(.7)^k$, 1^k y $(-1)^k$, respectivamente.

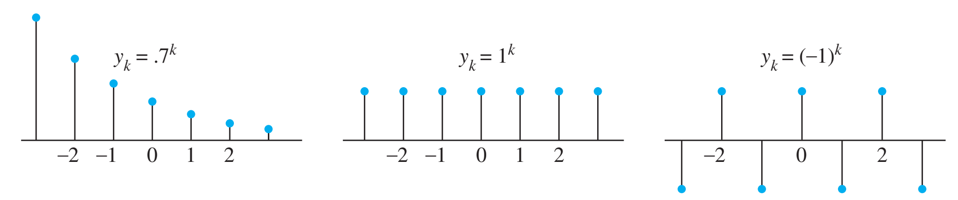


FIGURA 1 Tres señales en $\mathbb{S}$.

Es evidente que las señales digitales surgen en las ingenierías eléctrica y de sistemas de control, pero también se generan sucesiones de datos discretos en biología, física, economía, demografía, y muchas otras áreas, dondequiera que un proceso se mida, o *muestree*, a intervalos de tiempo discretos. Cuando un proceso se inicia en un momento específico, algunas veces conviene escribir una señal como una sucesión de la forma $(y_0, y_1, y_2, \ldots)$. Los términos y_k para $k < 0$ se suponen cero o simplemente se omiten.

EJEMPLO 1 Los cristalinos sonidos de un reproductor de discos compactos se producen a partir de música que se ha muestreado a razón de 44,100 veces por segundo. Vea la figura 2. En cada medición, la amplitud de la señal de música se registra como un número, por ejemplo, y_k. La música original está compuesta por muchos sonidos diferentes de diversas frecuencias, pero la sucesión $\{y_k\}$ contiene suficiente información para reproducir todas las frecuencias del sonido hasta los 20,000 ciclos por segundo, más de lo que puede percibir el oído humano.

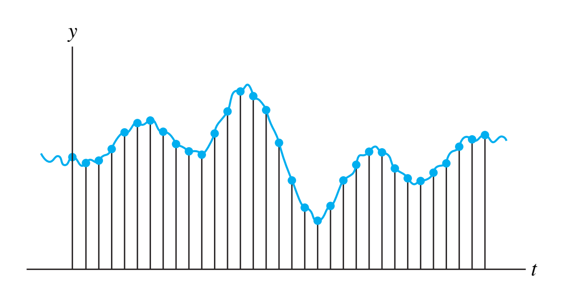


FIGURA 2 Muestra de datos de una señal musical.

Independencia lineal en el espacio $\mathbb{S}$ de señales

Para simplificar la notación, se considerará un conjunto de sólo tres señales en $\mathbb{S}$, por ejemplo $\{u_k\}$, $\{v_k\}$ y $\{w_k\}$. Éstas son linealmente independientes precisamente cuando la ecuación

$$c_1u_k + c_2v_k + c_3w_k = 0 \quad \text{para toda } k \tag{1}$$

implica que $c_1 = c_2 = c_3 = 0$. La frase "para toda k" significa para todos los enteros —positivos, negativos y cero—. También se podrían considerar señales que comenzaran con $k = 0$, por ejemplo, en cuyo caso "para toda k" significaría para todos los enteros $k \geq 0$.

Suponga que c_1, c_2, c_3 satisfacen (1). Entonces la ecuación en (1) se cumple para cualesquiera tres valores consecutivos de k, por ejemplo k, $k + 1$ y $k + 2$. Por lo tanto, (1) implica que

$$c_1u_{k+1} + c_2v_{k+1} + c_3w_{k+1} = 0 \quad \text{para toda } k$$

y

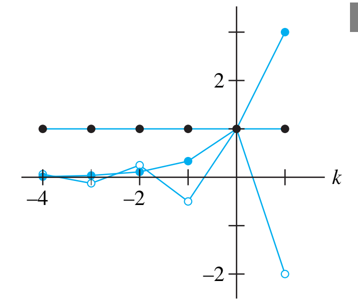
$$c_1u_{k+2} + c_2v_{k+2} + c_3w_{k+2} = 0 \quad \text{para toda } k$$

Por lo tanto, c_1, c_2, c_3 satisfacen

$$\begin{bmatrix} u_k & v_k & w_k \\ u_{k+1} & v_{k+1} & w_{k+1} \\ u_{k+2} & v_{k+2} & w_{k+2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{para toda } k \tag{2}$$

SG La prueba Casorati 4 a 31 (The Casorati Test 4-31)

La matriz de coeficientes incluida en este sistema se denomina **matriz de Casorati** de las señales, y el determinante de la matriz se conoce como el **casoratiano** de $\{u_k\}$, $\{v_k\}$ y $\{w_k\}$. Si la matriz de Casorati es invertible por lo menos para un valor de k, entonces (2) implicará que $c_1 = c_2 = c_3 = 0$, lo cual demostraría que las tres señales son linealmente independientes.



Las señales 1^k, $(-2)^k$ y 3^k.

EJEMPLO 2 Verifique si 1^k, $(-2)^k$ y 3^k son señales linealmente independientes.

Solución La matriz de Casorati es

$$\begin{bmatrix} 1^k & (-2)^k & 3^k \\ 1^{k+1} & (-2)^{k+1} & 3^{k+1} \\ 1^{k+2} & (-2)^{k+2} & 3^{k+2} \end{bmatrix}$$

Mediante operaciones por fila se puede mostrar muy fácilmente que esta matriz siempre es invertible. Sin embargo, es más rápido sustituir un valor para k —por ejemplo, $k = 0$— y reducir por filas la matriz numérica:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 3 \\ 1 & 4 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 3 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 0 & 10 \end{bmatrix}$$

La matriz de Casorati es invertible para $k = 0$. Por lo tanto, 1^k, $(-2)^k$ y 3^k son linealmente independientes.

Si una matriz de Casorati no es invertible, las señales asociadas que se están probando pueden ser o no linealmente dependientes. (Vea el ejercicio 33.) Sin embargo, puede mostrarse que si todas las *señales* son soluciones de la misma ecuación en diferencias homogénea (descrita a continuación), entonces la matriz de Casorati es invertible para

toda k y las señales son linealmente independientes o bien la matriz de Casorati no es invertible para toda k y las señales son linealmente dependientes. En la *guía de estudio* (*Study Guide*) se presenta una demostración utilizando transformaciones lineales.

Ecuaciones lineales en diferencias

Dados los escalares $a_0, \ldots, a_n$ con a_0 y a_n distintos de cero, y dada una señal $\{z_k\}$, la ecuación

$$a_0 y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{3}$$

se llama **ecuación lineal en diferencias** (o **relación de recurrencia lineal**) **de orden** $\boldsymbol{n}$. Por simplicidad, usualmente se toma a_0 como igual a 1. Si $\{z_k\}$ es la sucesión cero, la ecuación es **homogénea**; en caso contrario, la ecuación es **no homogénea**.

EJEMPLO 3 En el procesamiento digital de señales, una ecuación en diferencias, como la (3) anterior, describe un **filtro lineal**, y $a_0, \ldots, a_n$ se denominan **coeficientes de filtro**. Si $\{y_k\}$ se trata como la entrada y $\{z_k\}$ como la salida, entonces las soluciones de la ecuación homogénea asociada son las señales que se *eliminan* por filtración y se transforman en la señal cero. A continuación se alimentarán dos señales diferentes al filtro

$$.35y_{k+2} + .5y_{k+1} + .35y_k = z_k$$

Aquí .35 es una abreviatura para $\sqrt{2}/4$. La primera señal se crea muestreando la señal continua $y = \cos(\pi t/4)$ en valores enteros de t, como en la figura 3(a). La señal discreta es

$$\{y_k\} = \{\ldots, \cos(0), \cos(\pi/4), \cos(2\pi/4), \cos(3\pi/4), \ldots\}$$

Por simplicidad, se escribirá $\pm .7$ en lugar de $\pm\sqrt{2}/2$, de manera que

$$\{y_k\} = \{\ldots, \underset{\substack{\uparrow \\ k=0}}{1}, .7, 0, -.7, -1, -.7, 0, .7, 1, .7, 0, \ldots\}$$

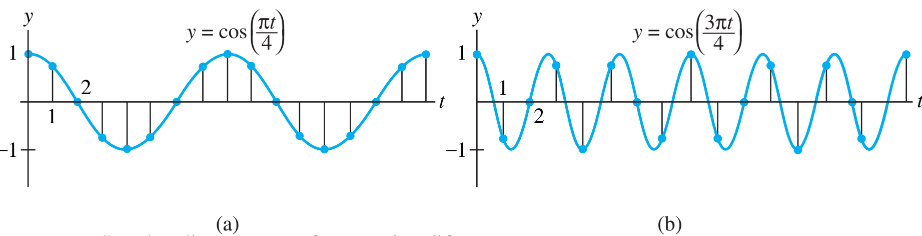


FIGURA 3 Señales discretas con frecuencias diferentes.

En la tabla 1 se muestra un cálculo de la sucesión de salida $\{z_k\}$, donde .35(.7) es una abreviatura para $(\sqrt{2}/4)(\sqrt{2}/2) = .25$. La salida es $\{y_k\}$, desplazada un término.

TABLA 1 Cálculo de la salida de un filtro

k	y_k	y_{k+1}	y_{k+2}	$.35y_k$	$+ .5y_{k+1}$	$+ .35y_{k+2}$	$=$	z_k
0	1	.7	0	.35(1)	+ .5(.7)	+ .35(0)	=	.7
1	.7	0	−.7	.35(.7)	+ .5(0)	+ .35(−.7)	=	0
2	0	−.7	−1	.35(0)	+ .5(−.7)	+ .35(−1)	=	−.7
3	−.7	−1	−.7	.35(−.7)	+ .5(−1)	+ .35(−.7)	=	−1
4	−1	−.7	0	.35(−1)	+ .5(−.7)	+ .35(0)	=	−.7
5	−.7	0	.7	.35(−.7)	+ .5(0)	+ .35(.7)	=	0
⋮	⋮							⋮

Se produce una señal de entrada diferente a partir de la señal de mayor frecuencia $y = \cos(3\pi t/4)$, la cual se muestra en la figura 3(b). Muestreando con la misma rapidez que antes se produce una nueva sucesión de entrada:

$$\{w_k\} = \{\ldots, \underset{\substack{\uparrow \\ k=0}}{1}, -.7, 0, .7, -1, .7, 0, -.7, 1, -.7, 0, \ldots\}$$

Cuando $\{w_k\}$ se alimenta al filtro, la salida es la sucesión cero. El filtro, llamado *filtro pasabajos*, deja pasar $\{y_k\}$, pero detiene la frecuencia mayor $\{w_k\}$.

En muchas aplicaciones, se especifica una sucesión $\{z_k\}$ para el miembro derecho de una ecuación en diferencias (3), y una $\{y_k\}$ que satisface (3) se denomina **solución** de la ecuación. El siguiente ejemplo muestra cómo encontrar soluciones para una ecuación homogénea.

EJEMPLO 4 Es frecuente que las soluciones de una ecuación homogénea en diferencias sean de la forma $y_k = r^k$ para alguna r. Encuentre algunas soluciones para la ecuación

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k \tag{4}$$

Solución Sustituya r^k por y_k en la ecuación y factorice el miembro izquierdo:

$$r^{k+3} - 2r^{k+2} - 5r^{k+1} + 6r^k = 0 \tag{5}$$

$$r^k(r^3 - 2r^2 - 5r + 6) = 0$$

$$r^k(r-1)(r+2)(r-3) = 0 \tag{6}$$

Como (5) es equivalente a (6), r^k satisface la ecuación en diferencias (4) si, y sólo si, r satisface (6). Entonces 1^k, $(-2)^k$, y 3^k son todas soluciones de (4). Por ejemplo, para verificar que 3^k es una solución de (4), calcule

$$\begin{aligned} &3^{k+3} - 2\cdot 3^{k+2} - 5\cdot 3^{k+1} + 6\cdot 3^k \\ &= 3^k(27 - 18 - 15 + 6) = 0 \quad \text{para toda } k \end{aligned}$$

En general, una señal distinta de cero r^k satisface la ecuación en diferencias

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1}y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

si, y sólo si, r es una raíz de la **ecuación auxiliar**

$$r^n + a_1 r^{n-1} + \cdots + a_{n-1} r + a_n = 0$$

No se considerará el caso en que r es una raíz repetida de la ecuación auxiliar. Cuando la ecuación auxiliar tiene una *raíz compleja*, la ecuación en diferencias presenta soluciones de la forma $s^k \cos k\omega$ y $s^k \operatorname{sen} k\omega$, para s y ω constantes. Esto ocurrió en el ejemplo 3.

Conjuntos solución de ecuaciones lineales en diferencias

Dadas $a_1, \ldots, a_n$, considere la función $T : \mathbb{S} \to \mathbb{S}$ que transforma una señal $\{y_k\}$ en una señal $\{w_k\}$ determinada mediante

$$w_k = y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k$$

Se comprueba de inmediato que T es una transformación *lineal.* Esto implica que el conjunto solución de la ecuación homogénea

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

es el núcleo de T (el conjunto de señales que T mapea en la señal cero) y, por lo tanto, el conjunto solución es un *subespacio* de $\mathbb{S}$. Cualquier combinación lineal de soluciones es de nuevo una solución.

El teorema siguiente, simple pero básico, proporcionará más información acerca de los conjuntos solución de las ecuaciones en diferencias.

TEOREMA 16

Si $a_n \neq 0$ y si $\{z_k\}$ está dada, la ecuación

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{7}$$

tiene una solución única para cualesquiera $y_0, \ldots, y_{n-1}$ especificadas.

DEMOSTRACIÓN Si se especifican $y_0, \ldots, y_{n-1}$, use (7) para *definir*

$$y_n = z_0 - [\, a_1 y_{n-1} + \cdots + a_{n-1} y_1 + a_n y_0 \,]$$

Y ahora que $y_1, \ldots, y_n$ están especificadas, use (7) para definir y_{n+1}. En general, use la relación de recurrencia

$$y_{n+k} = z_k - [\, a_1 y_{k+n-1} + \cdots + a_n y_k \,] \tag{8}$$

para definir y_{n+k} para $k \geq 0$. Para definir y_k para $k < 0$, use la relación de recurrencia

$$y_k = \frac{1}{a_n} z_k - \frac{1}{a_n} [\, y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} \,] \tag{9}$$

Esto produce una señal que satisface (7). Recíprocamente, cualquier señal que satisface (7) para toda k desde luego que satisface (8) y (9), así que la solución de (7) es única. ■

TEOREMA 17

El conjunto H de todas las soluciones de la ecuación lineal en diferencias homogénea de grado n

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k \tag{10}$$

es un espacio vectorial de dimensión n.

DEMOSTRACIÓN Como se explicó con anterioridad, H es un subespacio de $\mathbb{S}$ porque H es el núcleo de una transformación lineal. Para $\{y_k\}$ en H, sea $F\{y_k\}$ el vector en $\mathbb{R}^n$ dado por $(y_0, y_1, \ldots, y_{n-1})$. Se comprueba inmediatamente que $F : H \to \mathbb{R}^n$ es una transformación lineal. Dado cualquier vector $(y_0, y_1, \ldots, y_{n-1})$ en $\mathbb{R}^n$, el teorema 16 establece que hay una única señal $\{y_k\}$ en H tal que $F\{y_k\} = (y_0, y_1, \ldots, y_{n-1})$. Esto significa que F es una transformación lineal uno a uno de H sobre $\mathbb{R}^n$; es decir, F es un isomorfismo. Entonces $\dim H = \dim \mathbb{R}^n = n$. (Vea el ejercicio 32 de la sección 4.5.) ■

EJEMPLO 5 Encuentre una base para el conjunto de todas las soluciones de la ecuación en diferencias

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k$$

Solución En este momento el trabajo realizado en álgebra lineal será muy útil. Se sabe, por los ejemplos 2 y 4, que 1^k, $(-2)^k$, y 3^k son soluciones linealmente independientes. En general, puede resultar difícil comprobar directamente que un conjunto de señales *genera* el espacio solución. Pero aquí no hay problema debido a dos teoremas clave —el teorema 17, el cual muestra que la dimensión del espacio solución es exactamente tres, y el teorema de la base de la sección 4.5, el cual establece que un conjunto linealmente independiente con n vectores en un espacio vectorial n-dimensional es automáticamente una base—. Así, 1^k, $(-2)^k$, y 3^k forman una base para el espacio solución. ■

La manera estándar de describir la "solución general" de (10) es presentar una base para el subespacio de todas las soluciones. Una base así comúnmente se denomina **conjunto fundamental de soluciones** de (10). En la práctica, si se encuentran n señales linealmente independientes que satisfagan (10), generarán automáticamente la dimensión n del espacio solución, como se explicó en el ejemplo 5.

Ecuaciones no homogéneas

La solución general de la ecuación en diferencias no homogénea

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{11}$$

puede escribirse como una solución específica de (11), más una combinación lineal arbitraria de un conjunto fundamental de soluciones de la ecuación homogénea (10) correspondiente. Este hecho es análogo al resultado de la sección 1.5 acerca de cómo los conjuntos solución de $A\mathbf{x} = \mathbf{b}$ y de $A\mathbf{x} = \mathbf{0}$ son paralelos. Ambos resultados tienen la misma explicación: la función $\mathbf{x} \mapsto A\mathbf{x}$ es lineal, y la función que transforma la señal $\{y_k\}$ en la señal $\{z_k\}$ en (11) es lineal. Vea el ejercicio 35.

EJEMPLO 6 Compruebe que la señal $y_k = k^2$ satisface la ecuación en diferencias

$$y_{k+2} - 4y_{k+1} + 3y_k = -4k \quad \text{para toda } k \tag{12}$$

Después encuentre una descripción de todas las soluciones de esta ecuación.

Solución Sustituya y_k por k^2 en el miembro izquierdo de (12):

$$\begin{aligned}(k+2)^2 - 4(k+1)^2 + 3k^2 \\ &= (k^2+4k+4) - 4(k^2+2k+1) + 3k^2 \\ &= -4k\end{aligned}$$

Así que k^2 es, en efecto, una solución de (12). El siguiente paso es resolver la ecuación homogénea

$$y_{k+2} - 4y_{k+1} + 3y_k = 0 \tag{13}$$

La ecuación auxiliar es

$$r^2 - 4r + 3 = (r-1)(r-3) = 0$$

Las raíces son $r = 1, 3$. Entonces, dos soluciones de la ecuación en diferencias homogénea son 1^k y 3^k. Desde luego, éstas no son múltiplos una de la otra, así que son señales linealmente independientes. De acuerdo con el teorema 17, el espacio solución es bidimensional, así que 3^k y 1^k forman una base para el conjunto de soluciones de (13). Al trasladar ese conjunto mediante una solución específica de la ecuación no homogénea (12), se obtiene la solución general de (12):

$$k^2 + c_1 1^k + c_2 3^k, \quad \text{or} \quad k^2 + c_1 + c_2 3^k$$

En la figura 4 se proporciona una visualización geométrica de los dos conjuntos solución. Cada punto de la figura corresponde a una señal en $\mathbb{S}$.

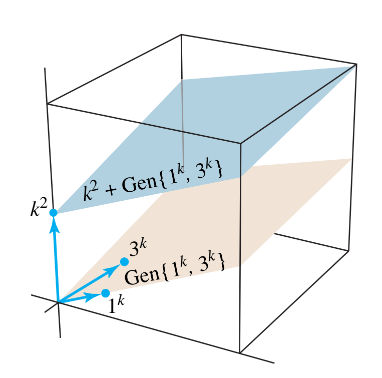


FIGURA 4
Conjuntos solución de las ecuaciones en diferencias (12) y (13).

Reducción a sistemas de ecuaciones de primer orden

Una manera moderna de estudiar una ecuación en diferencias lineal homogénea de orden n es reemplazarla por un sistema equivalente de ecuaciones en diferencias de primer orden, escrito en la forma

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{para toda } k$$

donde los vectores $\mathbf{x}_k$ están en $\mathbb{R}^n$ y A es una matriz de $n \times n$.

Ya se estudió un ejemplo sencillo de esta ecuación en diferencias (con valores vectoriales) en la sección 1.10. En las secciones 4.9 y 5.6 se darán otros ejemplos.

EJEMPLO 7 Escriba la siguiente ecuación en diferencias como un sistema de primer orden:

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k$$

Solución Para cada k, establezca

$$\mathbf{x}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

La ecuación en diferencias estipula que $y_{k+3} = -6y_k + 5y_{k+1} + 2y_{k+2}$, entonces

$$\mathbf{x}_{k+1} = \begin{bmatrix} y_{k+1} \\ y_{k+2} \\ y_{k+3} \end{bmatrix} = \begin{bmatrix} 0 & + & y_{k+1} & + & 0 \\ 0 & + & 0 & + & y_{k+2} \\ -6y_k & + & 5y_{k+1} & + & 2y_{k+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix} \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

Esto es,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{para toda } k, \quad \text{donde } A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix}$$

En general, la ecuación

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

puede reescribirse como $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para toda k, donde

$$\mathbf{x}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+n-1} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

Lecturas adicionales

Hamming, R. W., *Digital Filters*, 2a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 1983), págs. 1-37.

Kelly, W. G., y A. C. Peterson, *Difference Equations*, 2a. ed. (San Diego: Harcourt-Academic Press, 2001).

Mickens, R. E., *Difference Equations*, 2a. ed. (Nueva York: Van Nostrand Reinhold, 1990), págs. 88-141.

Oppenheim, A. V., y A. S. Willsky, *Signals and Systems*, 2a. ed. (Upper Saddle River, NJ: Prentice-Hall, 1997), págs. 1-14, 21-30, 38-43.

PROBLEMA DE PRÁCTICA

Puede mostrarse que las señales 2^k, $3^k \operatorname{sen} \frac{k\pi}{2}$, y $3^k \cos \frac{k\pi}{2}$ son soluciones de

$$y_{k+3} - 2y_{k+2} + 9y_{k+1} - 18y_k = 0$$

Muestre que estas señales forman una base para el conjunto de todas las soluciones de la ecuación en diferencias.

4.8 EJERCICIOS

Compruebe que las señales de los ejercicios 1 y 2 son soluciones de la ecuación en diferencias dada.

1. $2^k, (-4)^k; y_{k+2} + 2y_{k+1} - 8y_k = 0$

2. $3^k, (-3)^k; y_{k+2} - 9y_k = 0$

Muestre que las señales de los ejercicios 3 a 6 forman una base para el conjunto de soluciones de la ecuación en diferencias dada.

3. Las señales y ecuación del ejercicio 1.

4. Las señales y ecuación del ejercicio 2.

5. $(-3)^k, k(-3)^k; y_{k+2} + 6y_{k+1} + 9y_k = 0$

6. $5^k \cos\frac{k\pi}{2}, 5^k \operatorname{sen}\frac{k\pi}{2};\ y_{k+2} + 25y_k = 0$

En los ejercicios 7 a 12, suponga que las señales enlistadas son soluciones de la ecuación en diferencias dada. Determine si las señales forman una base para el espacio solución de la ecuación. Justifique sus respuestas usando los teoremas adecuados.

7. $1^k, 2^k, (-2)^k;\ y_{k+3} - y_{k+2} - 4y_{k+1} + 4y_k = 0$

8. $2^k, 4^k, (-5)^k;\ y_{k+3} - y_{k+2} - 22y_{k+1} + 40y_k = 0$

9. $1^k, 3^k \cos\frac{k\pi}{2}, 3^k \operatorname{sen}\frac{k\pi}{2};\ y_{k+3} - y_{k+2} + 9y_{k+1} - 9y_k = 0$

10. $(-1)^k, k(-1)^k, 5^k;\ y_{k+3} - 3y_{k+2} - 9y_{k+1} - 5y_k = 0$

11. $(-1)^k, 3^k;\ y_{k+3} + y_{k+2} - 9y_{k+1} - 9y_k = 0$

12. $1^k, (-1)^k;\ y_{k+4} - 2y_{k+2} + y_k = 0$

En los ejercicios 13 a 16, encuentre una base para el espacio solución de las ecuaciones en diferencias. Demuestre que las soluciones encontradas generan el conjunto solución.

13. $y_{k+2} - y_{k+1} + \frac{2}{9}y_k = 0$

14. $y_{k+2} - 7y_{k+1} + 12y_k = 0$

15. $y_{k+2} - 25y_k = 0$

16. $16y_{k+2} + 8y_{k+1} - 3y_k = 0$

Los ejercicios 17 y 18 se refieren a un modelo sencillo de la economía nacional descrito mediante la ecuación en diferencias:

$$Y_{k+2} - a(1+b)Y_{k+1} + abY_k = 1 \tag{14}$$

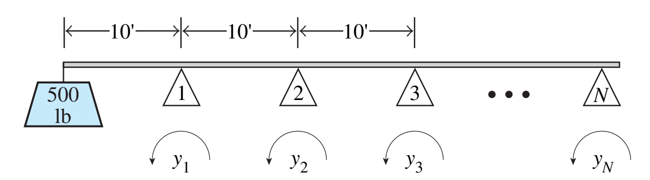
Aquí Y_k es el ingreso nacional total durante el año k, a es una constante menor que 1, llamada *propensión marginal al consumo*, y b es una *constante de ajuste* positiva que describe cómo los cambios en los gastos del consumidor afectan la tasa anual de inversión privada.[1]

17. Encuentre la solución general de (14) cuando $a = .9$ y $b = \frac{4}{9}$. ¿Qué le sucede a Y_k cuando k crece? [*Sugerencia:* Encuentre primero una solución específica de la forma $Y_k = T$, donde T es una constante, llamada nivel de equilibrio del ingreso nacional.]

18. Encuentre la solución general de (14) cuando $a = .9$ y $b = .5$.

Una viga ligera en voladizo está sostenida en N puntos que tienen una separación de 10 pies, y se coloca un peso de 500 lb en el extremo de la viga, a 10 pies del primer soporte, como indica la figura. Sea y_k el momento flector en el k-ésimo soporte. Entonces $y_1 = 5000$ pies por libra. Suponga que la viga está sujeta rígidamente al N-ésimo soporte y que el momento flector allí es cero. En los puntos intermedios, los momentos satisfacen la *ecuación de tres momentos*

$$y_{k+2} + 4y_{k+1} + y_k = 0 \quad \text{para } k = 1, 2, \ldots, N-2 \tag{15}$$



Momentos flectores en una viga en voladizo.

19. Encuentre la solución general de la ecuación en diferencias (15). Justifique su respuesta.

20. Encuentre la solución específica de (15) que satisface las *condiciones de frontera* $y_1 = 5000$ y $y_N = 0$. (En la respuesta aparece N.)

21. Cuando se produce una señal a partir de una sucesión de mediciones realizadas durante un proceso (una reacción química, un flujo de calor a través de un tubo, el movimiento del brazo de un robot, etc.), frecuentemente dicha señal contiene *ruido* aleatorio producido por errores de medición. Un método estándar de preprocesamiento de datos para reducir el ruido consiste en suavizar o filtrar los datos. Un filtro sencillo es un *promedio móvil* que reemplaza cada y_k por su promedio con los dos valores adyacentes:

$$\tfrac{1}{3}y_{k+1} + \tfrac{1}{3}y_k + \tfrac{1}{3}y_{k-1} = z_k \quad \text{para } k = 1, 2, \ldots$$

Suponga que una señal y_k, para $k = 0, \ldots, 14$, es

9, 5, 7, 3, 2, 4, 6, 5, 7, 6, 8, 10, 9, 5, 7

Use un filtro para calcular $z_1, \ldots, z_{13}$. Trace una gráfica de línea quebrada que superponga las señales original y suavizada.

22. Sea $\{y_k\}$ la secuencia producida al muestrear la señal continua $2\cos\frac{\pi t}{4} + \cos\frac{3\pi t}{4}$ en $t = 0, 1, 2, \ldots$, como indica la figura. Los valores de y_k, empezando con $k = 0$, son

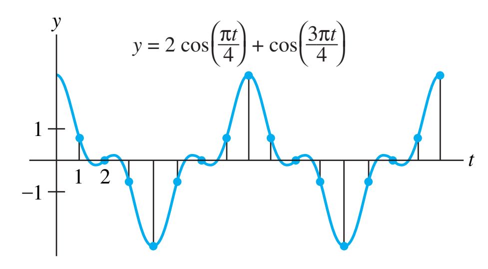
3, .7, 0, −.7, −3, −.7, 0, .7, 3, .7, 0, . . .

donde .7 es una abreviatura para $\sqrt{2}/2$.

a. Encuentre la señal de salida $\{z_k\}$ cuando $\{y_k\}$ se alimenta al filtro del ejemplo 3.

b. Explique cómo y por qué la salida encontrada en el inciso (a) se relaciona con los cálculos del ejemplo 3.

[1] Por ejemplo, vea *Discrete Dynamical Systems*, por James T. Sandefur, (Oxford: Clarendon Press, 1990), págs. 267-276. El *modelo acelerador multiplicador* original se debe al economista P. A. Samuelson.



Los datos muestreados de $2\cos\frac{\pi t}{4} + \cos\frac{3\pi t}{4}$.

Los ejercicios 23 y 24 se refieren a una ecuación en diferencias de la forma $y_{k+1} - ay_k = b$, para constantes adecuadas a y b.

23. Un préstamo de \$10,000 tiene una tasa de interés del 1% mensual y un pago de \$450 cada mes. El préstamo se hace en el mes $k = 0$ y el primer pago se realiza un mes después, en $k = 1$. Para $k = 0, 1, 2, \ldots$, sea y_k el saldo no pagado del préstamo después del k-ésimo pago mensual. Entonces

y_1	= 10,000	+ (.01)10,000	− 450
Nuevo saldo	Saldo actual	Interés agregado	Pago

a. Escriba una ecuación en diferencias que se cumpla para $\{y_k\}$.

b. **[M]** Elabore una tabla que muestre k y el saldo y_k en el mes k. Enliste el programa o las pulsaciones de tecla que haya usado para crear la tabla.

c. **[M]** ¿Cuál será el valor de k al efectuar el último pago? ¿De cuánto será el último pago? ¿Cuánto dinero en total habrá pagado el deudor?

24. En el tiempo $k = 0$, se efectúa una inversión inicial de \$1000 en una cuenta de ahorros que rinde un 6% de interés anual, compuesto mensualmente. (La tasa de interés por mes es de .005.) Cada mes posterior a la inversión inicial se agregan \$200 a la cuenta. Para $k = 0, 1, 2, \ldots$, sea y_k la cantidad a que asciende la cuenta en el momento k, justo después de haber hecho el depósito.

a. Escriba una ecuación en diferencias que se cumpla para $\{y_k\}$.

b. **[M]** Elabore una tabla que muestre k y la cantidad total habida en la cuenta de ahorros en el mes $k = 0$, para $k = 0$ hasta 60. Enliste su programa o las pulsaciones de tecla que necesitó para crear la tabla.

c. **[M]** ¿Cuánto habrá en la cuenta después de dos años (esto es, 24 meses), cuatro años, y cinco años? ¿Cuánto del saldo será por intereses a los cinco años?

En los ejercicios 25 a 28, muestre que la señal dada es una solución de la ecuación en diferencias. Luego encuentre la solución general de la ecuación en diferencias.

25. $y_k = k^2;\ \ y_{k+2} + 3y_{k+1} - 4y_k = 10k + 7$

26. $y_k = 1 + k;\ \ y_{k+2} - 8y_{k+1} + 15y_k = 8k + 2$

27. $y_k = 2 - 2k;\ \ y_{k+2} - \frac{9}{2}y_{k+1} + 2y_k = 3k + 2$

28. $y_k = 2k - 4;\ \ y_{k+2} + \frac{3}{2}y_{k+1} - y_k = 1 + 3k$

Escriba las ecuaciones en diferencias para los ejercicios 29 y 30 como sistemas de primer orden, $\mathbf{x}_{k+1} = A\mathbf{x}_k$, para toda k.

29. $y_{k+4} - 6y_{k+3} + 8y_{k+2} + 6y_{k+1} - 9y_k = 0$

30. $y_{k+3} - \frac{3}{4}y_{k+2} + \frac{1}{16}y_k = 0$

31. ¿La siguiente ecuación en diferencias es de orden 3? Explique su respuesta.

$$y_{k+3} + 5y_{k+2} + 6y_{k+1} = 0$$

32. ¿Cuál es el orden de la siguiente ecuación en diferencias? Explique su respuesta.

$$y_{k+3} + a_1y_{k+2} + a_2y_{k+1} + a_3y_k = 0$$

33. Sean $y_k = k^2$ y $z_k = 2k|k|$. ¿Las señales $\{y_k\}$ y $\{z_k\}$ son linealmente independientes? Evalúe la matriz de Casorati asociada $C(k)$ para $k = 0$, $k = -1$, y $k = -2$, y analice los resultados.

34. Sean f, g, h funciones linealmente independientes definidas para todos los números reales, y construya tres señales tomando muestras de los valores de las funciones en los enteros:

$$u_k = f(k), \qquad v_k = g(k), \qquad w_k = h(k)$$

¿Las señales deben ser linealmente independientes en $\mathbb{S}$? Analice el planteamiento.

35. Sean a y b dos números distintos de cero. Demuestre que la función T definida por $T\{y_k\} = \{w_k\}$, donde

$$w_k = y_{k+2} + ay_{k+1} + by_k$$

es una transformación lineal de $\mathbb{S}$ en $\mathbb{S}$.

36. Sean V un espacio vectorial y $T : V \to V$ una transformación lineal. Dado $\mathbf{z}$ en V, suponga que $\mathbf{x}_p$ en V cumple $T(\mathbf{x}_p) = \mathbf{z}$, y sea $\mathbf{u}$ cualquier vector en el núcleo de T. Muestre que $\mathbf{u} + \mathbf{x}_p$ satisface la ecuación no homogénea $T(\mathbf{x}) = \mathbf{z}$.

37. Sea $\mathbb{S}_0$ el vector espacial de todas las sucesiones de la forma $(y_0, y_1, y_2, \ldots)$, y defina transformaciones lineales T y D de $\mathbb{S}_0$ en $\mathbb{S}_0$ mediante

$$T(y_0, y_1, y_2, \ldots) = (y_1, y_2, y_3, \ldots)$$
$$D(y_0, y_1, y_2, \ldots) = (0, y_0, y_1, y_2, \ldots)$$

Muestre que $TD = I$ (la transformación identidad sobre $\mathbb{S}_0$) y que aún $DT \neq I$.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Examine la matriz de Casorati:

$$C(k) = \begin{bmatrix} 2^k & 3^k \operatorname{sen} \frac{k\pi}{2} & 3^k \cos \frac{k\pi}{2} \\ 2^{k+1} & 3^{k+1} \operatorname{sen} \frac{(k+1)\pi}{2} & 3^{k+1} \cos \frac{(k+1)\pi}{2} \\ 2^{k+2} & 3^{k+2} \operatorname{sen} \frac{(k+2)\pi}{2} & 3^{k+2} \cos \frac{(k+2)\pi}{2} \end{bmatrix}$$

Asuma $k = 0$ y reduzca por filas la matriz para comprobar que tiene tres posiciones pivote y, por lo tanto, es invertible:

$$C(0) = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 0 \\ 4 & 0 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -2 \\ 0 & 0 & -13 \end{bmatrix}$$

La matriz de Casorati es invertible en $k = 0$, así que las señales son linealmente independientes. Puesto que hay tres señales y el espacio solución H de la ecuación en diferencias tiene dimensión 3 (teorema 17), las señales forman una base para H, de acuerdo con el teorema de la base.

4.9 APLICACIONES A CADENAS DE MARKOV

Las cadenas de Markov que se describen en esta sección se usan como modelos matemáticos de una amplia variedad de situaciones en biología, química, ingeniería, física, los negocios y otros campos. En cada caso, el modelo se usa para describir un experimento o medición que se realiza varias veces de la misma forma, donde el resultado de cada ensayo del experimento será una de varias posibilidades, y donde el resultado de un ensayo depende solamente del ensayo inmediato anterior.

Por ejemplo, si la población de una ciudad y sus suburbios se midiera cada año, entonces un vector como

$$\mathbf{x}_0 = \begin{bmatrix} .60 \\ .40 \end{bmatrix} \tag{1}$$

podría indicar que el 60% de la población vive en la ciudad y el 40% en los suburbios. Los decimales en $\mathbf{x}_0$ suman 1 porque dan cuenta de la población total de la región. Los porcentajes son más convenientes para los propósitos de este ejercicio que los totales de población.

Un vector con entradas no negativas que suman 1 se llama **vector de probabilidad**. Una **matriz estocástica** es una matriz cuadrada cuyas columnas son vectores de probabilidad. Una **cadena de Markov** es una sucesión de vectores de probabilidad $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$, . . . , junto con una matriz estocástica P, tal que

$$\mathbf{x}_1 = P\mathbf{x}_0, \quad \mathbf{x}_2 = P\mathbf{x}_1, \quad \mathbf{x}_3 = P\mathbf{x}_2, \quad \ldots$$

Entonces la cadena de Markov se describe mediante la ecuación en diferencias de primer orden

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{para } k = 0, 1, 2, \ldots$$

Cuando una cadena de Markov de vectores en $\mathbb{R}^n$ describe un sistema o una sucesión de experimentos, las entradas en $\mathbf{x}_k$ enumeran, respectivamente, las probabilidades

de que el sistema esté en cada uno de n estados posibles o que el resultado de un experimento sea uno de los n posibles resultados. Por esta razón, frecuentemente se llama a $\mathbf{x}_k$ **vector de estado**.

EJEMPLO 1 En la sección 1.10 se examinó un modelo del movimiento de la población entre una ciudad y sus suburbios. Vea la figura 1. La migración anual entre estas dos partes de la región metropolitana estaba gobernada por la *matriz de migración M*:

$$\begin{array}{cccl} & \multicolumn{2}{c}{\text{De:}} & \\ & \text{Ciudad} & \text{Suburbios} & \text{A:} \\ M = & \left[\begin{array}{c} .95 \\ .05 \end{array}\right. & \left.\begin{array}{c} .03 \\ .97 \end{array}\right] & \begin{array}{l} \text{Ciudad} \\ \text{Suburbios} \end{array} \end{array}$$

Esto es, cada año el 5% de la población de la ciudad se muda a los suburbios y el 3% de la población de los suburbios se muda a la ciudad. Las columnas de M son vectores de probabilidad, así que M es una matriz estocástica. Suponga que la población de la región en el año 2000 es de 600,000 habitantes en la ciudad y 400,000 en los suburbios. Entonces la distribución inicial de la población en la región está dada por $\mathbf{x}_0$ en (1). ¿Cuál es la distribución de la población en 2001? ¿En 2002?

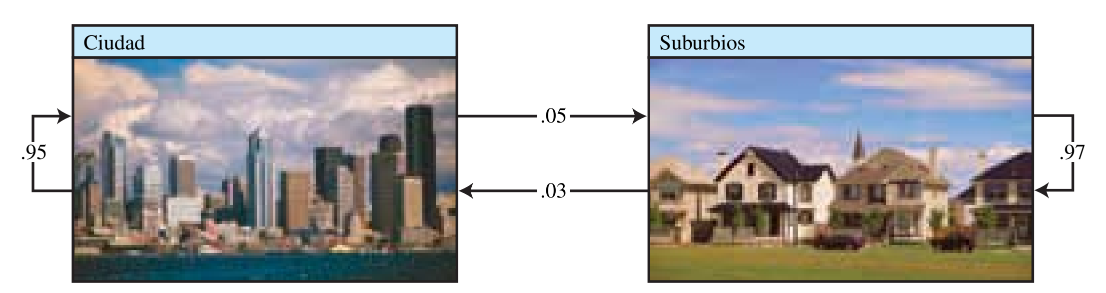


FIGURA 1 Porcentaje anual de migración entre la ciudad y los suburbios.

Solución En el ejemplo 3 de la sección 1.10 se vio que, después de un año, el vector de población $\begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$ cambió a

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

Al dividir ambos lados de esta ecuación entre la población total de un millón, y utilizar el hecho de que $kM\mathbf{x} = M(k\mathbf{x})$, se tiene

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} .600 \\ .400 \end{bmatrix} = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$$

El vector $\mathbf{x}_1 = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$ proporciona la distribución de población en 2001. Esto es, el 58.2% de la población de la región vivía en la ciudad y el 41.8% vivía en los suburbios.

De manera similar, la distribución de la población en 2002 se describe mediante un vector $\mathbf{x}_2$, donde

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} .582 \\ .418 \end{bmatrix} = \begin{bmatrix} .565 \\ .435 \end{bmatrix}$$

EJEMPLO 2 Suponga que los resultados de la votación en una elección al congreso estadounidense en cierto distrito electoral están representados mediante un vector $\mathbf{x}$ en $\mathbb{R}^3$:

$$\mathbf{x} = \begin{bmatrix} \text{porcentaje que vota por los demócratas (D)} \\ \text{porcentaje que vota por los republicanos (R)} \\ \text{porcentaje que vota por los libertarios (L)} \end{bmatrix}$$

Suponga que se registran los resultados de la elección al congreso cada dos años mediante un vector de este tipo y que el resultado de una elección depende solamente de los resultados de la elección anterior. Entonces la sucesión de vectores que describe los votos cada dos años puede ser una cadena de Markov. Como ejemplo de matriz estocástica P para esta cadena, se toma

$$\begin{array}{c} \\ \\ P = \end{array} \begin{array}{c} \text{De:} \\ \begin{array}{ccc} \text{D} & \text{R} & \text{L} \end{array} \\ \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \end{array} \begin{array}{c} \\ \text{A:} \\ \text{D} \\ \text{R} \\ \text{L} \end{array}$$

Las entradas incluidas en la primera columna, etiquetada como D, describen lo que las personas que votan por demócratas en una elección harán en la siguiente elección. Aquí se ha supuesto que el 70% de las personas votará D nuevamente, el 20% votará R, y un 10% votará L. Para las demás columnas de P se proporciona una interpretación similar. En la figura 2 se presenta un diagrama para esta matriz.

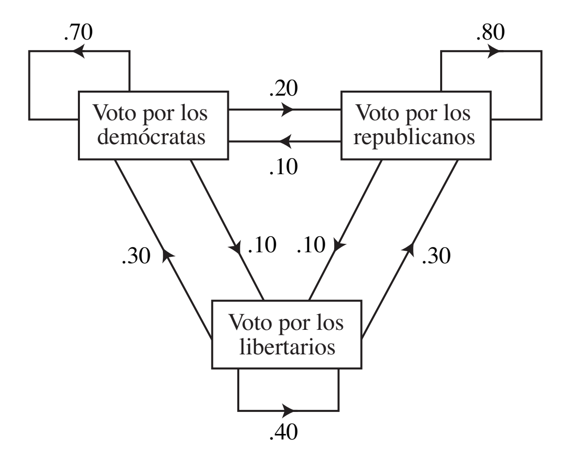


FIGURA 2 Cambios de preferencia de una elección a la siguiente.

Si los porcentajes de "transición" permanecen constantes durante muchos años, de una elección a la siguiente, entonces la sucesión de vectores que proporcionan los resul-

tados de las votaciones forma una cadena de Markov. Suponga que en una elección los resultados están dados por

$$\mathbf{x}_0 = \begin{bmatrix} .55 \\ .40 \\ .05 \end{bmatrix}$$

Determine el resultado probable para la siguiente elección y para la elección sucesiva.

Solución Los resultados de la siguiente elección se describen mediante el vector de estado $\mathbf{x}_1$ y los de la elección posterior por medio de $\mathbf{x}_2$, donde

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \begin{bmatrix} .55 \\ .40 \\ .05 \end{bmatrix} = \begin{bmatrix} .440 \\ .445 \\ .115 \end{bmatrix} \quad \begin{matrix} \text{44\% votarán D.} \\ \text{44.5\% votarán R.} \\ \text{11.5\% votarán L.} \end{matrix}$$

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \begin{bmatrix} .440 \\ .445 \\ .115 \end{bmatrix} = \begin{bmatrix} .3870 \\ .4785 \\ .1345 \end{bmatrix} \quad \begin{matrix} \text{38.7\% votarán D.} \\ \text{47.8\% votarán R.} \\ \text{13.5\% votarán L.} \end{matrix}$$

Para entender por qué $\mathbf{x}_1$ proporciona realmente los resultados de la siguiente elección, suponga que 1000 personas votaron en la "primera" elección, con 550 votantes a favor de D, 400 a favor de R, y 50 a favor de L. (Vea los porcentajes en $\mathbf{x}_0$.) En la siguiente elección, el 70% de los 550 votará D de nuevo, el 10% de los 400 cambiará de R a D, y el 30% de los 50 cambiará de L a D. Entonces el total de votos para D será

$$.70(550) + .10(400) + .30(50) = 385 + 40 + 15 = 440 \tag{2}$$

Así que el 44% de los votos en la próxima elección será para el candidato D. El cálculo en (2) es, esencialmente, el mismo que se usó para determinar la primera entrada de $\mathbf{x}_1$. Pueden realizarse cálculos análogos para las otras entradas de $\mathbf{x}_1$, para las entradas de $\mathbf{x}_2$, y así sucesivamente.

Predicción del futuro lejano

El aspecto más interesante de las cadenas de Markov es el estudio del comportamiento de una cadena a largo plazo. Veamos, en el ejemplo 2, ¿qué puede decirse acerca de los votos luego de que han pasado muchas elecciones (si se supone que la matriz estocástica dada sigue describiendo los porcentajes de transición de una elección a la siguiente)? O, ¿qué le sucede a la distribución de población del ejemplo 1 "a la larga"? Antes de contestar estas preguntas, se presenta un ejemplo numérico.

EJEMPLO 3 Sea $P = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Considere un sistema cuyo estado se describe mediante la cadena de Markov $\mathbf{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, \ldots$. ¿Qué le sucede al sistema con el paso del tiempo? Para encontrar la respuesta, determine los vectores de estado $\mathbf{x}_1, \ldots, \mathbf{x}_{15}$.

Solución

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix}$$

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}\begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix} = \begin{bmatrix} .37 \\ .45 \\ .18 \end{bmatrix}$$

$$\mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}\begin{bmatrix} .37 \\ .45 \\ .18 \end{bmatrix} = \begin{bmatrix} .329 \\ .525 \\ .146 \end{bmatrix}$$

Los resultados de los cálculos posteriores se presentan enseguida, con entradas redondeadas a cuatro o cinco cifras significativas.

$$\mathbf{x}_4 = \begin{bmatrix} .3133 \\ .5625 \\ .1242 \end{bmatrix}, \quad \mathbf{x}_5 = \begin{bmatrix} .3064 \\ .5813 \\ .1123 \end{bmatrix}, \quad \mathbf{x}_6 = \begin{bmatrix} .3032 \\ .5906 \\ .1062 \end{bmatrix}, \quad \mathbf{x}_7 = \begin{bmatrix} .3016 \\ .5953 \\ .1031 \end{bmatrix}$$

$$\mathbf{x}_8 = \begin{bmatrix} .3008 \\ .5977 \\ .1016 \end{bmatrix}, \quad \mathbf{x}_9 = \begin{bmatrix} .3004 \\ .5988 \\ .1008 \end{bmatrix}, \quad \mathbf{x}_{10} = \begin{bmatrix} .3002 \\ .5994 \\ .1004 \end{bmatrix}, \quad \mathbf{x}_{11} = \begin{bmatrix} .3001 \\ .5997 \\ .1002 \end{bmatrix}$$

$$\mathbf{x}_{12} = \begin{bmatrix} .30005 \\ .59985 \\ .10010 \end{bmatrix}, \quad \mathbf{x}_{13} = \begin{bmatrix} .30002 \\ .59993 \\ .10005 \end{bmatrix}, \quad \mathbf{x}_{14} = \begin{bmatrix} .30001 \\ .59996 \\ .10002 \end{bmatrix}, \quad \mathbf{x}_{15} = \begin{bmatrix} .30001 \\ .59998 \\ .10001 \end{bmatrix}$$

Estos vectores parecen tender a $\mathbf{q} = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$. Las probabilidades apenas cambian de un valor de k al próximo. Observe que el cálculo siguiente es exacto (sin error de redondeo):

$$P\mathbf{q} = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}\begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix} = \begin{bmatrix} .15 + .12 + .03 \\ .09 + .48 + .03 \\ .06 + 0 \ + .04 \end{bmatrix} = \begin{bmatrix} .30 \\ .60 \\ .10 \end{bmatrix} = \mathbf{q}$$

Cuando el sistema está en estado $\mathbf{q}$, no hay cambio en el sistema de una medición a la siguiente.

Vectores de estado estacionario

Si P es una matriz estocástica, entonces un **vector de estado estacionario** (o **vector de equilibrio**) para P es un vector de probabilidad $\mathbf{q}$ tal que

$$P\mathbf{q} = \mathbf{q}$$

Puede mostrarse que cada matriz estocástica tiene un vector de estado estacionario. En el ejemplo 3, $\mathbf{q}$ es un vector de estado estacionario para P.

EJEMPLO 4 El vector de probabilidad $\mathbf{q} = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$ es un vector de estado estacionario para la matriz de migración de población M dada en el ejemplo 1, porque

$$M\mathbf{q} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .375 \\ .625 \end{bmatrix} = \begin{bmatrix} .35625 + .01875 \\ .01875 + .60625 \end{bmatrix} = \begin{bmatrix} .375 \\ .625 \end{bmatrix} = \mathbf{q}$$

Si la población total de la región metropolitana del ejemplo 1 es de un millón de habitantes, entonces la **q** del ejemplo 4 correspondería a tener 375,000 personas en la ciudad y 625,000 en los suburbios. Al final de un año, la migración *desde* la ciudad sería de (.05)(375,000) = 18,750 personas, y la migración *a* la ciudad desde los suburbios sería de (.03)(625,000) = 18,750 personas. En consecuencia, la población en la ciudad no cambiaría. Por lo mismo, la población suburbana sería estable.

En el siguiente ejemplo se muestra cómo *encontrar* un vector de estado estacionario.

EJEMPLO 5 Sea $P = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$. Encuentre un vector de estado estacionario para P.

Solución Primero, resuelva la ecuación $P\mathbf{x} = \mathbf{x}$.

$$\begin{aligned} P\mathbf{x} - \mathbf{x} &= \mathbf{0} \\ P\mathbf{x} - I\mathbf{x} &= \mathbf{0} \qquad \text{Recuerde, de la sección 1.4, que } I\mathbf{x} = \mathbf{x}. \\ (P - I)\mathbf{x} &= \mathbf{0} \end{aligned}$$

Para la P mencionada,

$$P - I = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -.4 & .3 \\ .4 & -.3 \end{bmatrix}$$

Para encontrar las soluciones de $(P - I)\mathbf{x} = \mathbf{0}$, reduzca por filas la matriz aumentada:

$$\begin{bmatrix} -.4 & .3 & 0 \\ .4 & -.3 & 0 \end{bmatrix} \sim \begin{bmatrix} -.4 & .3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3/4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Entonces $x_1 = \frac{3}{4}x_2$ y x_2 es libre. La solución general es $x_2 \begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$.

Enseguida, elija una base sencilla para el espacio solución. Una selección evidente es $\begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$, pero una mejor elección sin fracciones es $\mathbf{w} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ (la cual corresponde a $x_2 = 4$).

Por último, encuentre un vector de probabilidad en el conjunto de todas las soluciones de $P\mathbf{x} = \mathbf{x}$. Este proceso es fácil, puesto que cada solución es un múltiplo de la **w** anterior. Divida **w** entre la suma de sus entradas y obtenga

$$\mathbf{q} = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$$

Como comprobación, calcule

$$P\mathbf{q} = \begin{bmatrix} 6/10 & 3/10 \\ 4/10 & 7/10 \end{bmatrix} \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix} = \begin{bmatrix} 18/70 + 12/70 \\ 12/70 + 28/70 \end{bmatrix} = \begin{bmatrix} 30/70 \\ 40/70 \end{bmatrix} = \mathbf{q}$$

El teorema siguiente muestra que lo sucedido en el ejemplo 3 es típico de muchas matrices estocásticas. Se afirma que una matriz estocástica es **regular** si alguna potencia de la matriz P^k contiene sólo entradas estrictamente positivas. Para la P del ejemplo 3, se tiene

$$P^2 = \begin{bmatrix} .37 & .26 & .33 \\ .45 & .70 & .45 \\ .18 & .04 & .22 \end{bmatrix}$$

Como toda entrada de P^2 es estrictamente positiva, P es una matriz estocástica regular.

Además, una sucesión de vectores $\{\mathbf{x}_k : k = 1, 2, \ldots\}$ **converge** hacia un vector $\mathbf{q}$ cuando $k \to \infty$ si las entradas de las $\mathbf{x}_k$ pueden volverse tan cercanas como se desee a las entradas correspondientes de $\mathbf{q}$ tomando una k lo suficientemente grande.

TEOREMA 18

Si P es una matriz de $n \times n$ estocástica regular, entonces P tiene un único vector de estado estacionario $\mathbf{q}$. Además, si $\mathbf{x}_0$ es cualquier estado inicial y $\mathbf{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, 2, \ldots$, entonces la cadena de Markov $\{\mathbf{x}_k\}$ converge hacia $\mathbf{q}$ cuando $k \to \infty$.

Este teorema se demuestra en textos estándar sobre cadenas de Markov. La parte sorprendente del teorema es que el estado inicial no tiene efecto sobre el comportamiento a largo plazo de la cadena de Markov. Posteriormente se verá (en la sección 5.2) por qué este hecho es cierto para varias de las matrices estocásticas que se estudian aquí.

EJEMPLO 6 En el ejemplo 2, ¿qué porcentaje de los electores es probable que vote por el candidato republicano en alguna elección celebrada dentro de muchos años, suponiendo que los resultados de las elecciones forman una cadena de Markov?

Solución Si se realizan cálculos a mano, el enfoque *erróneo* es elegir algún vector inicial $\mathbf{x}_0$ y calcular $\mathbf{x}_1, \ldots, \mathbf{x}_k$ para algún valor grande de k. No se tiene manera de saber cuántos vectores habrá que calcular, y no se puede estar seguro de los valores límite para las entradas de los $\mathbf{x}_k$.

El enfoque correcto consiste en calcular el vector de estado estacionario y aplicar el teorema 18. Dada P como en el ejemplo 2, forme $P - I$ restando 1 a cada entrada de la diagonal de P. Después reduzca por filas la matriz aumentada:

$$[\,(P - I) \quad \mathbf{0}\,] = \begin{bmatrix} -.3 & .1 & .3 & 0 \\ .2 & -.2 & .3 & 0 \\ .1 & .1 & -.6 & 0 \end{bmatrix}$$

Recuerde del trabajo realizado con decimales que la aritmética se simplifica al multiplicar cada fila por 10.[1]

$$\begin{bmatrix} -3 & 1 & 3 & 0 \\ 2 & -2 & 3 & 0 \\ 1 & 1 & -6 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -9/4 & 0 \\ 0 & 1 & -15/4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general de $(P - I)\mathbf{x} = \mathbf{0}$ es $x_1 = \frac{9}{4}x_3$, $x_2 = \frac{15}{4}x_3$, y x_3 es libre. Al elegir $x_3 = 4$ se obtiene una base para el espacio solución cuyas entradas son enteros, y a partir de aquí se obtiene fácilmente el vector de estado estacionario cuyas entradas suman 1:

$$\mathbf{w} = \begin{bmatrix} 9 \\ 15 \\ 4 \end{bmatrix}, \quad \text{y} \quad \mathbf{q} = \begin{bmatrix} 9/28 \\ 15/28 \\ 4/28 \end{bmatrix} \approx \begin{bmatrix} .32 \\ .54 \\ .14 \end{bmatrix}$$

Las entradas de **q** describen la distribución de votos en una elección que se efectuará dentro de muchos años (si se asume que la matriz estocástica continúa describiendo los cambios de una elección a la siguiente). Así, en algún momento, alrededor del 54% de los votos será para el candidato republicano.

NOTA NUMÉRICA

Quizá el lector haya notado que si $\mathbf{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, \ldots$, entonces

$$\mathbf{x}_2 = P\mathbf{x}_1 = P(P\mathbf{x}_0) = P^2\mathbf{x}_0,$$

y, en general,

$$\mathbf{x}_k = P^k\mathbf{x}_0 \quad \text{para } k = 0, 1, \ldots$$

Para calcular un vector específico como $\mathbf{x}_3$, se necesitarán menos operaciones aritméticas si se determinan $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ en vez de calcular P^3 y $P^3\mathbf{x}_0$. Sin embargo, cuando P es pequeña —por ejemplo de 30×30— el tiempo de máquina necesario para realizar los cálculos resulta insignificante con ambos métodos, y se podría preferir un comando para calcular $P^3\mathbf{x}_0$ porque requiere menos golpes de tecla por parte del usuario.

PROBLEMAS DE PRÁCTICA

1. Suponga que los residentes de una región metropolitana se mudan de acuerdo con las probabilidades dadas en la matriz de migración del ejemplo 1, y que se elige un residente "al azar". Entonces un vector de estado para cierto año puede interpretarse como algo que indica la probabilidad de que la persona sea residente de la ciudad o de los suburbios en ese momento.

a. Suponga que la persona seleccionada es en este momento un residente de la ciudad, así que $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. ¿Cuál es la probabilidad de que la persona viva en los suburbios el próximo año?

[1] *Advertencia:* No multiplique sólo P por 10. En vez de esto, multiplique por 10 la matriz aumentada para la ecuación $(P - I)\mathbf{x} = \mathbf{0}$.

b. ¿Cuál es la probabilidad de que la persona esté viviendo en los suburbios dentro de dos años?

2. Sean $P = \begin{bmatrix} .6 & .2 \\ .4 & .8 \end{bmatrix}$ y $\mathbf{q} = \begin{bmatrix} .3 \\ .7 \end{bmatrix}$. ¿Es **q** un vector de estado estacionario para P?

3. ¿Qué porcentaje de la población del ejemplo 1 vivirá en los suburbios dentro de muchos años?

4.9 EJERCICIOS

1. Una aldea remota recibe señales de radio de dos estaciones, una estación de noticias y otra de música. De los radioescuchas que sintonizan la estación de noticias, el 70% seguirá escuchándola después de una interrupción de la estación que ocurre cada media hora, mientras que el 30% cambiará a la estación de música en el momento de la interrupción. De los radioescuchas que sintonizan la estación de música, el 60% cambiará a la estación de noticias después de la interrupción de la estación, mientras que el 40% seguirá escuchando música. Suponga que a las 8:15 A.M. todos están oyendo las noticias.

a. Proporcione la matriz estocástica que describe cómo los radioescuchas tienden a cambiar de estación en cada interrupción. Etiquete las filas y las columnas.

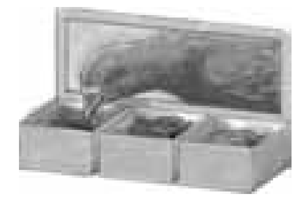
b. Proporcione el vector de estado inicial.

c. ¿Qué porcentaje de los radioescuchas estará oyendo música a las 9:25 A.M. (después de las interrupciones de estación de las 8:30 y 9:00 A.M.)?

2. Un animal de laboratorio puede comer cualquiera de tres alimentos cada día. Los registros del laboratorio muestran que si el animal elige un alimento en un ensayo, hay una probabilidad del 50% de que prefiera el mismo alimento en el siguiente ensayo, y elegirá los otros alimentos con probabilidades iguales del 25%.

a. ¿Cuál es la matriz estocástica para esta situación?

b. Si el animal elige el alimento #1 en un ensayo inicial, ¿cuál es la probabilidad de que prefiera el alimento #2 en el segundo ensayo después del ensayo inicial?



3. En un día determinado, un estudiante está sano o enfermo. De los estudiantes que están sanos hoy, el 95% lo estará mañana. De los que están enfermos hoy, el 55% seguirá enfermo mañana.

a. ¿Cuál es la matriz estocástica para esta situación?

b. Suponga que el lunes el 20% de los estudiantes está enfermo. ¿Qué fracción o porcentaje de los estudiantes es probable que siga enfermo el martes? ¿El miércoles?

c. Si un estudiante está bien hoy, ¿cuál es la probabilidad de que siga bien dentro de dos días?

4. El clima en Columbus puede ser bueno, regular o malo en un día determinado. Si el tiempo es bueno hoy, hay una probabilidad del 60% de que mañana sea bueno, una probabilidad del 30% de que sea regular, y una probabilidad del 10% de que sea malo. Si el clima de hoy es regular, será bueno mañana con una probabilidad de .40 y regular con una probabilidad de .30. Por último, si el tiempo es malo hoy, será bueno mañana con una probabilidad de .40 y regular con una probabilidad de .50.

a. ¿Cuál es la matriz estocástica para esta situación?

b. Suponga que hoy se tiene una probabilidad del 50% de buen clima y una probabilidad del 50% de clima regular. ¿Cuál es la probabilidad de que mañana el clima sea malo?

c. Suponga que la predicción del clima para el lunes es de 40% para clima regular y de 60% para mal clima. ¿Cuál es la probabilidad de tener buen clima el miércoles?

En los ejercicios 5 a 8, encuentre el vector de estado estacionario

5. $\begin{bmatrix} .1 & .6 \\ .9 & .4 \end{bmatrix}$

6. $\begin{bmatrix} .8 & .5 \\ .2 & .5 \end{bmatrix}$

7. $\begin{bmatrix} .7 & .1 & .1 \\ .2 & .8 & .2 \\ .1 & .1 & .7 \end{bmatrix}$

8. $\begin{bmatrix} .7 & .2 & .2 \\ 0 & .2 & .4 \\ .3 & .6 & .4 \end{bmatrix}$

9. Determine si $P = \begin{bmatrix} .2 & 1 \\ .8 & 0 \end{bmatrix}$ es una matriz estocástica regular.

10. Determine si $P = \begin{bmatrix} 1 & .2 \\ 0 & .8 \end{bmatrix}$ es una matriz estocástica regular.

11. a. Encuentre el vector de estado estacionario para la cadena de Markov del ejercicio 1.

b. En algún momento, ya avanzado el día, ¿qué fracción de los radioescuchas estará sintonizando las noticias?

12. Con referencia al ejercicio 2, ¿qué comida preferirá el animal después de muchos ensayos?

13. a. Encuentre el vector de estado estacionario para la cadena de Markov del ejercicio 3.

b. ¿Cuál es la probabilidad de que luego de muchos días cierto estudiante esté enfermo? ¿Importa si la persona está enferma hoy?

14. Con referencia al ejercicio 4, a la larga, ¿qué tan probable es que el clima de Columbus sea bueno en un día determinado?

15. **[M]** La Unidad de Investigación Demográfica del Departamento de Finanzas del Estado de California proporcionó datos para la siguiente matriz de migración, la cual describe el movimiento de la población dentro de Estados Unidos durante 1989. En 1989, cerca del 11.7% de la población total vivía en California. ¿Qué porcentaje de la población total vivirá eventualmente en California si las probabilidades de migración que se dan permanecen constantes durante muchos años?

De: CA	Resto de EUA	A:
.9821	.0029	California
.0179	.9971	Resto de EUA

16. **[M]** En Detroit, Hertz Rent A Car posee una flotilla de unos 2000 automóviles. El patrón de lugares de renta y sitios de retorno está dado en fracciones en la tabla siguiente. En un día típico, ¿cuántos coches estarán listos, aproximadamente, para rentarse en la ubicación del centro?

Automóviles rentados en:

Aeropuerto de la ciudad	Centro	Aeropuerto metropolitano	Devueltos en:
.90	.01	.09	Aeropuerto de la ciudad
.01	.90	.01	Centro
.09	.09	.90	Aeropuerto metropolitano

17. Sea P una matriz estocástica de $n \times n$. El siguiente argumento muestra que la ecuación $P\mathbf{x} = \mathbf{x}$ tiene una solución no trivial. (De hecho, existe una solución de estado estacionario con entradas no negativas. En algunos textos avanzados se da una demostración.) Justifique cada una de las siguientes afirmaciones. (Mencione un teorema cuando sea apropiado.)

a. Si todas las otras filas de $P - I$ se suman a la fila inferior, el resultado es una fila de ceros.

b. Las filas de $P - I$ son linealmente dependientes.

c. La dimensión del espacio fila de $P - I$ es menor que n.

d. $P - I$ tiene un espacio nulo no trivial.

18. Demuestre que toda matriz estocástica de 2×2 tiene por lo menos un vector de estado estacionario. Cualquier matriz de este tipo se escribe como $P = \begin{bmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{bmatrix}$, donde α y β son constantes entre 0 y 1. (Existen dos vectores de estado estacionario linealmente independientes si $\alpha = \beta = 0$. En caso contrario, sólo hay un vector.)

19. Sea S la matriz fila de $1 \times n$ con un 1 en cada columna, $S = [1 \ \ 1 \ \ \cdots \ \ 1]$

a. Explique por qué un vector $\mathbf{x}$ en $\mathbb{R}^n$ es un vector de probabilidad si, y sólo si, sus entradas son no negativas y $S\mathbf{x} = 1$. (Una matriz de 1×1 como el producto $S\mathbf{x}$ normalmente se escribe sin los corchetes matriciales.)

b. Sea P una matriz estocástica de $n \times n$. Explique por qué $SP = S$.

c. Sean P una matriz estocástica de $n \times n$ y $\mathbf{x}$ un vector de probabilidad. Demuestre que $P\mathbf{x}$ también es un vector de probabilidad.

20. Use el ejercicio 19 para demostrar que si P es una matriz estocástica de $n \times n$, entonces P^2 también lo es.

21. **[M]** Examine las potencias de una matriz estocástica regular.

a. Calcule P^k, para $k = 2, 3, 4, 5$, cuando

$$P = \begin{bmatrix} .3355 & .3682 & .3067 & .0389 \\ .2663 & .2723 & .3277 & .5451 \\ .1935 & .1502 & .1589 & .2395 \\ .2047 & .2093 & .2067 & .1765 \end{bmatrix}$$

Muestre los cálculos con cuatro cifras decimales. ¿Qué les sucede a las columnas de P^k cuando k crece? Calcule el vector de estado estacionario para P.

b. Calcule Q^k para $k = 10, 20, \ldots, 80$, cuando

$$Q = \begin{bmatrix} .97 & .05 & .10 \\ 0 & .90 & .05 \\ .03 & .05 & .85 \end{bmatrix}$$

(Para que Q^k se mantenga estable con cuatro cifras decimales, puede requerirse $k = 116$ o más.) Calcule el vector de estado estacionario para Q. Formule una conjetura acerca de lo que podría ser cierto para cualquier matriz estocástica regular.

c. Use el teorema 18 para explicar lo que se encontró en (a) y (b).

22. **[M]** Compare dos métodos para encontrar el vector de estado estacionario **q** para una matriz estocástica regular P: (1) calcular **q** igual que en el ejemplo 5, o (2) calcular P^k para algún valor grande de k y usar una de las columnas de P^k como una aproximación de **q**. [La *guía de estudio* (*Study Guide*) describe un programa de *base nula* que casi automatiza al método (1)].

Experimente con las matrices estocásticas aleatorias más grandes que permita su programa de matrices, y use $k = 100$ o algún otro valor grande. Por cada método, describa el tiempo necesario para pulsar las teclas y correr su programa. (Algunas versiones de MATLAB tienen comandos `flops` y `tic...toc` para registrar el número de operaciones de punto flotante y el tiempo total que emplea MATLAB.) Compare las ventajas de cada método y exprese cuál prefiere.

Soluciones a los problemas de práctica

1. a. Como el 5% de los residentes de la ciudad se mudará a los suburbios en un lapso de un año, hay una probabilidad del 5% de elegir a una persona que lo haga. Sin saber más acerca de la persona, se afirma que hay una probabilidad del 5% de que la persona se mude a los suburbios. Este hecho está contenido en la segunda entrada del vector de estado $\mathbf{x}_1$, donde

$$\mathbf{x}_1 = M\mathbf{x}_0 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} .95 \\ .05 \end{bmatrix}$$

b. La probabilidad de que la persona esté viviendo en los suburbios después de dos años es del 9.6%, porque

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .95 \\ .05 \end{bmatrix} = \begin{bmatrix} .904 \\ .096 \end{bmatrix}$$

2. El vector de estado estacionario satisface $P\mathbf{x} = \mathbf{x}$. Puesto que

$$P\mathbf{q} = \begin{bmatrix} .6 & .2 \\ .4 & .8 \end{bmatrix}\begin{bmatrix} .3 \\ .7 \end{bmatrix} = \begin{bmatrix} .32 \\ .68 \end{bmatrix} \neq \mathbf{q}$$

se concluye que **q** *no* es el vector de estado estacionario para P.

3. La M del ejemplo 1 es una matriz estocástica regular porque todas sus entradas son estrictamente positivas. Así, puede usarse el teorema 18. Ya se conoce el vector de estado estacionario del ejemplo 4. Entonces los vectores de distribución de la población $\mathbf{x}_k$ convergen hacia

$$\mathbf{q} = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$$

Con el tiempo, el 62.5% de la población vivirá en los suburbios.

CD Aplicaciones de cadenas de Markov (Applications of Markov Chains)

Capítulo 4 Ejercicios suplementarios

1. Señale cada enunciado como verdadero o falso. Justifique sus respuestas. (Si es verdadero, cite los hechos o teoremas adecuados. Si es falso, explique por qué o proporcione un contraejemplo para mostrar por qué el enunciado no es cierto en todos los casos.) En los incisos (a) a (f), $\mathbf{v}_1, \ldots, \mathbf{v}_p$ son vectores en un espacio vectorial de dimensión finita V, y $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

a. El conjunto de las combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$ es un espacio vectorial.

b. Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}\}$ genera V, entonces S genera V.

c. Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}\}$ es linealmente independiente, entonces S también lo es.

d. Si S es linealmente independiente, entonces S es una base para V.

e. Si Gen $S = V$, entonces algún subconjunto de S es una base para V.

f. Si dim $V = p$ y Gen $S = V$, entonces S no puede ser linealmente dependiente.

g. Un plano en $\mathbb{R}^3$ es un subespacio de dimensión 2.

h. Las columnas no pivote de una matriz son siempre linealmente dependientes.

i. Las operaciones por fila sobre una matriz A pueden cambiar las relaciones de dependencia lineal entre las filas de A.

j. Las operaciones por fila sobre una matriz pueden cambiar el espacio nulo.

k. El rango de una matriz es igual al número de filas distintas de cero.

l. Si una matriz A de $m \times n$ es equivalente por filas a una matriz escalonada U, y si U tiene k filas distintas de cero, entonces la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$ es $m - k$.

m. Si B se obtiene de una matriz A mediante varias operaciones elementales de fila, entonces rango B = rango A.

n. Las filas distintas de cero de una matriz A forman una base para Fil A.

o. Si las matrices A y B tienen la misma forma escalonada reducida, entonces Fil A = Fil B.

p. Si H es un subespacio de $\mathbb{R}^3$, entonces existe una matriz A de 3×3 tal que H = Col A.

q. Si A es de $m \times n$ y rango $A = m$, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno.

r. Si A es una matriz $m \times n$ y la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es suprayectiva, entonces rango $A = m$.

s. Una matriz de cambio de coordenadas siempre es invertible.

t. Si $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ son bases para un espacio vectorial V, entonces la j-ésima columna de la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es el vector de coordenadas $[\mathbf{c}_j]_{\mathcal{B}}$.

2. Encuentre una base para el conjunto de todos los vectores de la forma

$$\begin{bmatrix} a - 2b + 5c \\ 2a + 5b - 8c \\ -a - 4b + 7c \\ 3a + b + c \end{bmatrix}. \text{ (Sea cuidadoso.)}$$

3. Sea $\mathbf{u}_1 = \begin{bmatrix} -2 \\ 4 \\ -6 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, y W = Gen$\{\mathbf{u}_1, \mathbf{u}_2\}$. Encuentre una descripción *implícita* de W; esto es, determine un conjunto de una o más ecuaciones homogéneas que caractericen los puntos W. [*Pista:* ¿Cuándo está $\mathbf{b}$ en W?]

4. Explique lo que está mal en el siguiente análisis: Sea $\mathbf{f}(t) = 3 + t$ y $\mathbf{g}(t) = 3t + t^2$, y observe que $\mathbf{g}(t) = t\mathbf{f}(t)$. Entonces $\{\mathbf{f}, \mathbf{g}\}$ es linealmente dependiente porque $\mathbf{g}$ es un múltiplo de $\mathbf{f}$.

5. Considere los polinomios $\mathbf{p}_1(t) = 1 + t$, $\mathbf{p}_2(t) = 1 - t$, $\mathbf{p}_3(t) = 4$, $\mathbf{p}_4(t) = t + t^2$, y $\mathbf{p}_5(t) = 1 + 2t + t^2$, y sea H el subespacio $\mathbb{P}_5$ generado por el conjunto $S = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5\}$. Utilice el método descrito en la demostración del teorema del conjunto generador (sección 4.3) para producir una base de H. (Explique cómo seleccionar los elementos apropiados de S.)

6. Suponga que $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$ son polinomios específicos que generan un subespacio H bidimensional de $\mathbb{P}_5$. Describa cómo puede encontrarse una base para H examinando los cuatro polinomios y casi sin hacer cálculos.

7. ¿Qué tendría que saberse acerca del conjunto solución de un sistema de 18 ecuaciones lineales con 20 variables para asegurarse de que toda ecuación no homogénea asociada tiene solución? Analice el planteamiento.

8. Sea H un subespacio de dimensión n de un espacio vectorial V de dimensión n. Explique por qué $H = V$.

9. Sea $T: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal.

a. ¿Cuál es la dimensión del rango de T si T es una función uno a uno? Explique su respuesta.

b. ¿Cuál es la dimensión del núcleo de T (vea la sección 4.2) si T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$? Explique su respuesta.

10. Sea S un subconjunto linealmente independiente máximo de un espacio vectorial V. Esto es, S tiene la propiedad de que si un vector que no está en S se agrega a S, entonces el nuevo conjunto no seguirá siendo linealmente independiente. Demuestre que S debe ser una base para V. [*Pista:* ¿Qué pasaría si S fuera linealmente independiente pero no una base para V?]

11. Sea S un conjunto finito generador mínimo de un espacio vectorial V. Esto es, S tiene la propiedad de que si se elimina un vector de S, entonces el nuevo conjunto ya no genera V. Demuestre que S debe ser una base para V.

En los ejercicios 12 a 17 se desarrollan propiedades del rango que ocasionalmente son necesarias en algunas aplicaciones. Suponga que la matriz A es de $m \times n$.

12. Muestre a partir de (a) y (b) que el rango AB no puede exceder el rango de A o de B. (En general, el rango de un producto de matrices no excede el rango de ninguno de los factores.)

a. Muestre que si B es de $n \times p$, entonces rango $AB \leq$ rango A. [*Indicación:* Explique por qué todo vector del espacio columna de AB está en el espacio columna de A.]

b. Muestre que si B es de $n \times p$, entonces rango $AB \leq$ rango B. [*Sugerencia:* Use el inciso (a) para estudiar el rango $(AB)^T$.]

13. Muestre que si P es una matriz invertible de $m \times m$, entonces rango PA = rango A. [*Sugerencia:* Aplique el ejercicio 12 a PA y $P^{-1}(PA)$.]

14. Muestre que si Q es invertible, entonces rango $AQ =$ rango A [*Sugerencia:* Use el ejercicio 13 para estudiar rango$(AQ)^T$.]

15. Sea A una matriz de $m \times n$, y sea B una matriz de $n \times p$ tal que $AB = 0$. Muestre que rango A + rango $B \leq n$. [*Indicación:* Uno de los cuatro subespacios Nul A, Col A, Nul B y Col B está contenido en uno de los otros tres subespacios.]

16. Si A es una matriz de $m \times n$ de rango r, entonces una *factorización de rango* de A es una ecuación de la forma $A = CR$, donde C es una matriz de $m \times r$ de rango r y R es una matriz de $r \times n$ de rango r. Siempre existe una factorización de este tipo (ejercicio 38 en la sección 4.6). Dadas cualesquiera dos matrices de $m \times n$ A y B, use factorizaciones de rango de A y B para demostrar que

$$\text{rango}(A + B) \leq \text{rango } A + \text{rango } B$$

[*Sugerencia:* Escriba $A + B$ como el producto de dos matrices partidas.]

17. Una **submatriz** de una matriz A es cualquier matriz que resulta de eliminar algunas (o ninguna) filas y/o columnas de A. Puede demostrarse que A tiene rango r si, y sólo si, A contiene una submatriz invertible de $r \times r$ y ninguna submatriz cuadrada mayor es invertible. Demuestre parte de esta afirmación explicando (a) por qué una matriz de $m \times n$ de rango r tiene una submatriz A_1 de $m \times r$ de rango r, y (b) por qué A_1 tiene una submatriz invertible A_2 de $r \times r$.

El concepto de rango cumple un papel importante en ingeniería en los procesos de diseño de sistemas de control, como el del transbordador espacial mencionado en el ejemplo introductorio de este capítulo. Un *modelo en el espacio de estados* de un sistema de control incluye una ecuación en diferencias de la forma

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k \qquad \text{para } k = 0, 1, \ldots \tag{1}$$

donde A es de $n \times n$, B es de $n \times m$, $\{\mathbf{x}_k\}$ es una sucesión de “vectores de estado” en $\mathbb{R}^n$ que describen el estado del sistema en tiempos discretos, y $\{\mathbf{u}_k\}$ es una secuencia de *control*, o de *entrada*. Se afirma que el par (A, B) es **controlable si**

$$\text{rango } [B \quad AB \quad A^2B \quad \cdots \quad A^{n-1}B] = n \tag{2}$$

La matriz que aparece en (2) se denomina **matriz de controlabilidad** para el sistema. Si (A, B) es controlable, entonces el sistema puede controlarse, o conducirse desde el estado $\mathbf{0}$ hasta cualquier estado específico $\mathbf{v}$ (en $\mathbb{R}^n$) en, cuando mucho, n pasos, con sólo elegir una secuencia de control adecuada en $\mathbb{R}^m$. Este hecho se ilustra en el ejercicio 18 para $n = 4$ y $m = 2$. Un análisis más profundo de la controlabilidad puede consultarse en el sitio web de este texto (Estudio de caso para el capítulo 4).

WEB

18. Suponga que A es una matriz de 4×4 y B una matriz de 4×2, y considere que $\mathbf{u}_0, \ldots, \mathbf{u}_3$ representa una sucesión de vectores de entrada en $\mathbb{R}^2$.

a. Establezca $\mathbf{x}_0 = \mathbf{0}$, calcule $\mathbf{x}_1, \ldots, \mathbf{x}_4$ a partir de (1), y escriba una fórmula para $\mathbf{x}_4$ que involucre la matriz de controlabilidad M incluida en (2). (*Nota:* La matriz M se construye como una matriz partida. Aquí, su tamaño global es de 4×8.)

b. Suponga que (A, B) es controlable y que $\mathbf{v}$ es cualquier vector en $\mathbb{R}^4$. Explique por qué existe una secuencia de control $\mathbf{u}_0, \ldots, \mathbf{u}_3$ en $\mathbb{R}^2$ tal que $\mathbf{x}_4 = \mathbf{v}$.

Determine si los pares de matrices de los ejercicios 19 a 22 son controlables.

19. $A = \begin{bmatrix} .9 & 1 & 0 \\ 0 & -.9 & 0 \\ 0 & 0 & .5 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

20. $A = \begin{bmatrix} .8 & -.3 & 0 \\ .2 & .5 & 1 \\ 0 & 0 & -.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

21. **[M]** $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & -4.2 & -4.8 & -3.6 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

22. **[M]** $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -13 & -12.2 & -1.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

5

Valores propios y vectores propios

EJEMPLO INTRODUCTORIO

Sistemas dinámicos y los búhos manchados

En 1990, el búho manchado del norte se convirtió en el centro de una controversia nacional en Estados Unidos acerca del uso de los majestuosos bosques del Noroeste pacífico.

Los ecologistas convencieron al gobierno federal de que el búho manchado estaría condenado a la extinción si continuaba la tala en los bosques de crecimiento viejo (con árboles de más de 200 años), donde prefieren vivir los búhos. La industria maderera, anticipando la pérdida de entre 30,000 y 100,000 empleos como resultado de las nuevas restricciones a la tala por parte del gobierno, argumentó que el búho no debería clasificarse como una "especie en peligro de extinción" y citó varios informes científicos publicados para apoyar su caso.[1]

Atrapados en el fuego cruzado de los dos grupos en conflicto, los ecologistas matemáticos intensificaron sus esfuerzos por entender la dinámica poblacional del búho manchado. El ciclo de vida de un búho manchado se divide naturalmente en tres etapas: juvenil (hasta 1 año de edad), subadulto (1 a 2 años), y adulto (más de 2 años).

El búho se aparea de por vida durante las etapas de subadulto y adulto, empieza a reproducirse en la edad adulta, y vive hasta 20 años. Cada pareja de búhos requiere aproximadamente de 1000 hectáreas (4 millas cuadradas) como territorio base. Un momento crítico en el ciclo de vida es cuando los búhos jóvenes abandonan el nido. Para sobrevivir y convertirse en subadulto, un búho joven debe encontrar un nuevo territorio base (y generalmente una pareja).

Un primer paso para estudiar la dinámica poblacional es configurar un modelo de la población a intervalos anuales, en tiempos denotados mediante $k = 0, 1, 2, \ldots$. Por lo general, se supone que existe una relación 1:1 de machos a hembras en cada etapa de vida, y se cuentan exclusivamente las hembras. La población en el año k puede describirse por medio de un vector $\mathbf{x}_k = (j_k, s_k, a_k)$, donde j_k, s_k y a_k son las cantidades de hembras existentes en las etapas juvenil, subadulto y adulto, respectivamente.

[1]"The Great Spotted Owl War", *Reader's Digest*, noviembre de 1992, págs. 91-95.

Utilizando datos de campo de estudios demográficos reales, R. Lamberson y colaboradores consideraron el siguiente *modelo de matrices por etapas*:[2]

$$\begin{bmatrix} j_{k+1} \\ s_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \begin{bmatrix} j_k \\ s_k \\ a_k \end{bmatrix}$$

Aquí, la cantidad de nuevas hembras juveniles en el año $k+1$ es .33 veces la cantidad de hembras adultas en el año k (con base en el índice de natalidad promedio por pareja de búhos). También, el 18% de los jóvenes sobrevive para convertirse en subadultos, y un 71% de los subadultos y el 94% de los adultos sobreviven para ser contados como adultos.

El modelo de matrices por etapas es una ecuación en diferencias de la forma $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Una ecuación de este tipo se llama **sistema dinámico** (o **sistema dinámico lineal discreto**) porque describe los cambios experimentados en un sistema al paso del tiempo.

[2]R. H. Lamberson, R. McKelvey, B. R. Noon, y C. Voss, "A Dynamic Analysis of the Viability of the Northern Spotted Owl in a Fragmented Forest Environment", *Conservation Biology* **6** (1992), págs. 505-512. También, una comunicación privada del profesor Lamberson, 1993.

La tasa de supervivencia juvenil del 18% en la matriz por etapas de Lamberson es la entrada más afectada por la cantidad de bosque viejo disponible. De hecho, el 60% de los búhos juveniles normalmente sobrevive para dejar el nido, pero en la región de Willow Creek, California, estudiada por Lamberson y sus colegas, sólo el 30% de los jóvenes que dejaron el nido pudo encontrar nuevos territorios base; el resto pereció durante el proceso de búsqueda.

Una razón importante del fracaso de los búhos al tratar de encontrar nuevos territorios es el aumento en la fragmentación de las áreas con árboles viejos debido a la tala total de áreas diseminadas en los terrenos de crecimiento viejo. Cuando un búho deja la protección del bosque y cruza un área devastada, el riesgo de que sea atacado por un depredador aumenta de modo impresionante. En la sección 5.6 se mostrará que el modelo descrito anteriormente predice la eventual extinción del búho manchado, pero también que si el 50% de los búhos juveniles que sobreviven para dejar el nido encuentra nuevos territorios, entonces la población de búho manchado prosperará.

La meta de este capítulo es examinar detenidamente la acción de una transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ para obtener elementos que se visualicen fácilmente. Excepto por una breve digresión en la sección 5.4, todas las matrices del capítulo son cuadradas. Las principales aplicaciones descritas aquí son de sistemas dinámicos discretos, incluidos los de búhos manchados que se analizaron con anterioridad. Sin embargo, los conceptos básicos —vectores propios y valores propios— son útiles en todas las áreas de las matemáticas puras y aplicadas, y aparecen en contextos mucho más generales que los considerados aquí. Los valores propios también se usan para estudiar ecuaciones diferenciales y sistemas dinámicos *continuos*, proporcionan información crítica en el diseño de ingeniería, y se presentan naturalmente en campos como la física y la química.

5.1 VECTORES PROPIOS Y VALORES PROPIOS

Aunque una transformación $\mathbf{x} \mapsto A\mathbf{x}$ puede mover vectores en diversas direcciones, con frecuencia sucede que existen vectores especiales sobre los cuales la acción de A resulta muy sencilla.

EJEMPLO 1 Sean $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, y $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Las imágenes de $\mathbf{u}$ y $\mathbf{v}$ bajo la multiplicación por A se muestran en la figura 1. En realidad, $A\mathbf{v}$ es justamente $2\mathbf{v}$. Así que A sólo "estira" o dilata $\mathbf{v}$.

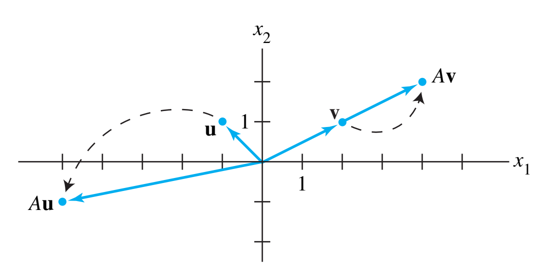


FIGURA 1 Efectos de la multiplicación por A.

Como ejemplo adicional, los lectores de la sección 4.9 recordarán que si A es una matriz estocástica, entonces el vector $\mathbf{q}$ de estado estacionario para A satisface la ecuación $A\mathbf{x} = \mathbf{x}$. Esto es, $A\mathbf{q} = 1 \cdot \mathbf{q}$.

En esta sección, se estudian ecuaciones como

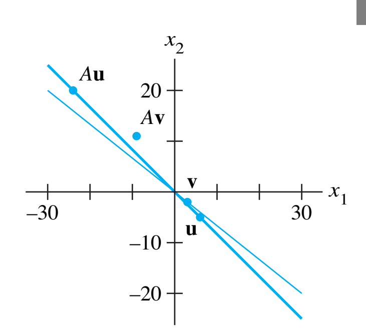
$$A\mathbf{x} = 2\mathbf{x} \quad \text{o bien} \quad A\mathbf{x} = -4\mathbf{x}$$

y se buscan vectores que sean transformados por A en múltiplos escalares de sí mismos.

DEFINICIÓN Un **vector propio** de una matriz A de $n \times n$ es un vector $\mathbf{x}$ diferente de cero tal que $A\mathbf{x} = \lambda\mathbf{x}$ para algún escalar λ. Un escalar λ se llama **valor propio** de A si existe una solución no trivial $\mathbf{x}$ de $A\mathbf{x} = \lambda\mathbf{x}$; una $\mathbf{x}$ como ésta se denomina *vector propio correspondiente a* λ.[1]

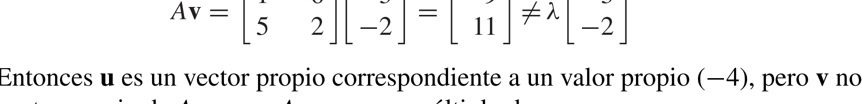
Es fácil determinar si un vector dado es un vector propio de una matriz. También resulta sencillo decidir si un escalar específico es un valor propio.

EJEMPLO 2 Sean $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, y $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. ¿Son $\mathbf{u}$ y $\mathbf{v}$ vectores propios de A?



$A\mathbf{u} = -4\mathbf{u}$, pero $A\mathbf{v} \neq \lambda\mathbf{v}$.

Solución



$$A\mathbf{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4\begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4\mathbf{u}$$

$$A\mathbf{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda\begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Entonces $\mathbf{u}$ es un vector propio correspondiente a un valor propio (-4), pero $\mathbf{v}$ no es un vector propio de A porque $A\mathbf{v}$ no es un múltiplo de $\mathbf{v}$.

EJEMPLO 3 Muestre que 7 es un valor propio de A en el ejemplo 2, y encuentre los vectores propios correspondientes.

Solución El escalar 7 es un valor propio de A si, y sólo si, la ecuación

$$A\mathbf{x} = 7\mathbf{x} \tag{1}$$

---

[1]Observe que, por definición, un vector propio debe ser *distinto de cero*, pero un valor propio sí puede ser cero. El caso en que el número 0 es un valor propio se analiza después del ejemplo 5.

tiene una solución no trivial. Pero (1) es equivalente a $A\mathbf{x} - 7\mathbf{x} = 0$, o bien

$$(A - 7I)\mathbf{x} = \mathbf{0} \tag{2}$$

Para resolver esta ecuación homogénea, forme la matriz

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

Desde luego, las columnas de $A - 7I$ son linealmente dependientes, así que (2) tiene soluciones no triviales. Entonces 7 *es* un valor propio de A. Para encontrar los vectores propios correspondientes, use operaciones por fila:

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

La solución general tiene la forma $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Cada vector de esta forma con $x_2 \neq 0$ es un vector propio correspondiente a $\lambda = 7$. ■

Advertencia: Aunque en el ejemplo 3 se utilizó reducción por filas para encontrar los *vectores propios*, ésta no puede usarse para calcular *valores propios*. Una forma escalonada de una matriz A generalmente no exhibe los valores propios de A.

La equivalencia de las ecuaciones (1) y (2) evidentemente se mantiene para cualquier λ que esté en lugar de $\lambda = 7$. Entonces λ es un valor propio de A si, y sólo si, la ecuación

$$(A - \lambda I)\mathbf{x} = \mathbf{0} \tag{3}$$

tiene una solución no trivial. El conjunto de *todas* las soluciones de (3) es justamente el espacio nulo de la matriz $A - \lambda I$. De manera que este conjunto es un *subespacio* de $\mathbb{R}^n$ y se llama el **espacio propio** de A correspondiente a λ. El espacio propio consiste en los vectores cero y en todos los vectores propios correspondientes a λ.

En el ejemplo 3 se muestra que para la A del ejemplo 2, el espacio propio correspondiente a $\lambda = 7$ consiste en *todos* los múltiplos de (1, 1), los cuales forman la línea que pasa por (1, 1) y el origen. Por el ejemplo 2, se puede comprobar que el espacio propio correspondiente a $\lambda = -4$ es la línea que pasa por (6, −5). Estos espacios propios se muestran en la figura 2, junto con los vectores propios (1, 1) y (3/2, −5/4) y la acción geométrica de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ sobre cada espacio propio.

EJEMPLO 4 Sea $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. Un valor propio de A es 2. Encuentre una base para el espacio propio correspondiente.

Solución Forme

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

y reduzca por filas la matriz aumentada para $(A - 2I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

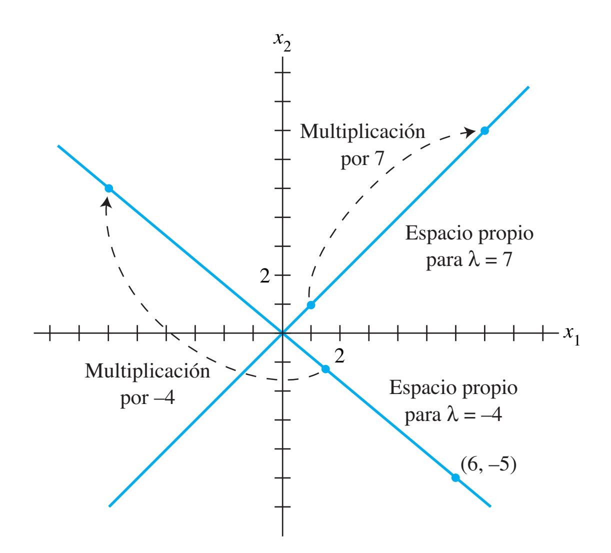


FIGURA 2 Espacios propios para $\lambda = -4$ y $\lambda = 7$.

En este punto se tiene la seguridad de que 2 sí es un valor propio de A porque la ecuación $(A - 2I)\mathbf{x} = \mathbf{0}$ tiene variables libres. La solución general es

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \quad x_2 \text{ y } x_3 \text{ son libres}$$

El espacio propio, mostrado en la figura 3, es un subespacio bidimensional de $\mathbb{R}^3$. Una base es

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

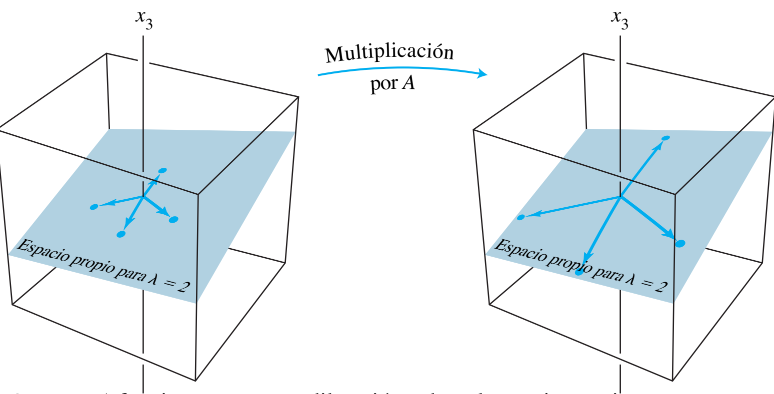


FIGURA 3 A funciona como una dilatación sobre el espacio propio.

NOTA NUMÉRICA

En el ejemplo 4 se muestra un buen método para efectuar cálculos manuales de vectores propios en los casos sencillos en que se conoce un valor propio. Por lo general, el uso de un programa de matrices y reducción por filas (con un valor propio específico) para encontrar un espacio propio funciona bien, pero no es totalmente confiable. El error de redondeo puede llevar ocasionalmente a una forma escalonada reducida con un número erróneo de pivotes. Los mejores programas de computadora calculan aproximaciones para los valores propios y los vectores propios simultáneamente, hasta el nivel de precisión deseado, para matrices que no sean muy grandes. El tamaño de las matrices que se puede analizar aumenta cada año conforme van mejorando la capacidad computacional y los programas de cómputo.

El teorema siguiente describe uno de los pocos casos especiales en que la determinación de valores propios puede hacerse con precisión. El cálculo de valores propios se analizará también en la sección 5.2.

TEOREMA 1 Los valores propios de una matriz triangular son las entradas de su diagonal principal.

DEMOSTRACIÓN En aras de la simplicidad, considere el caso de 3×3. Si A es triangular superior, entonces $A - \lambda I$ es de la forma

$$A - \lambda I = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix}$$

El escalar λ es un valor propio de A si, y sólo si, la ecuación $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial; esto es, si, y sólo si, la ecuación tiene una variable libre. Debido a las entradas cero en $A - \lambda I$, es fácil ver que $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tiene una variable libre si, y sólo si, por lo menos una de las entradas en la diagonal de $A - \lambda I$ es cero. Esto pasa si, y sólo si, λ es igual a alguna de las entradas a_{11}, a_{22}, a_{33} de A. Para el caso de que A sea triangular inferior, vea el ejercicio 28. ■

EJEMPLO 5 Sean $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$. Los valores propios de A son 3, 0 y 2. Los valores propios de B son 4 y 1. ■

¿Qué significa que una matriz A tenga un valor propio de 0, como en el ejemplo 5? Esto sucede si, y sólo si, la ecuación

$$A\mathbf{x} = 0\mathbf{x} \tag{4}$$

tiene una solución no trivial. Pero (4) es equivalente a $A\mathbf{x} = \mathbf{0}$, la cual tiene una solución no trivial si, y sólo si, A es no invertible. Entonces 0 *es un valor propio de A si, y sólo si, A es no invertible.* Este hecho se agregará al teorema de la matriz invertible presentado en la sección 5.2.

El teorema siguiente es fundamental y se necesitará más adelante. Su demostración ilustra un típico cálculo con vectores propios.

TEOREMA 2 Si $\mathbf{v}_1, \ldots, \mathbf{v}_r$ son vectores propios que corresponden a distintos valores propios $\lambda_1, \ldots, \lambda_r$ de una matriz A de $n \times n$, entonces el conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ es linealmente independiente.

DEMOSTRACIÓN Suponga que $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ es linealmente dependiente. Como $\mathbf{v}_1$ es distinto de cero, el teorema 7 de la sección 1.7 establece que uno de los vectores presente en el conjunto es una combinación lineal de los vectores precedentes. Sea p el índice mínimo tal que $\mathbf{v}_{p+1}$ es una combinación lineal de los vectores precedentes (linealmente independientes). Entonces existen escalares $c_1, \ldots, c_p$ tales que

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{v}_{p+1} \tag{5}$$

Si se multiplican ambos lados de (5) por A y se usa el hecho de que $A\mathbf{v}_k = \lambda_k\mathbf{v}_k$ para cada k, se obtiene

$$c_1A\mathbf{v}_1 + \cdots + c_pA\mathbf{v}_p = A\mathbf{v}_{p+1}$$
$$c_1\lambda_1\mathbf{v}_1 + \cdots + c_p\lambda_p\mathbf{v}_p = \lambda_{p+1}\mathbf{v}_{p+1} \tag{6}$$

Si se multiplican ambos lados de (5) por λ_{p+1} y se resta el resultado a (6), se tiene

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \cdots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0} \tag{7}$$

Como $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es linealmente independiente, los pesos en (7) son todos iguales a cero. Pero ninguno de los factores $\lambda_i - \lambda_{p+1}$, es cero, porque los valores propios son distintos. De aquí que $c_i = 0$ para $i = 1, \ldots, p$. Pero entonces (5) proclama $\mathbf{v}_{p+1} = \mathbf{0}$, lo cual es imposible. Por lo tanto, $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ no puede ser linealmente dependiente y, por lo tanto, debe ser linealmente independiente. ■

Vectores propios y ecuaciones en diferencias

Esta sección concluye mostrando cómo construir soluciones de la ecuación en diferencias de primer orden:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad (k = 0, 1, 2, \ldots) \tag{8}$$

Si A es una matriz $n \times n$, entonces (8) es una descripción *recursiva* de una sucesión $\{\mathbf{x}_k\}$ en $\mathbb{R}^n$. Una **solución** de (8) es una descripción explícita de $\{\mathbf{x}_k\}$, cuya fórmula para cada $\mathbf{x}_k$ no depende directamente de A ni de los términos precedentes en la sucesión excepto del término inicial $\mathbf{x}_0$.

La manera más simple de construir una solución para (8) es tomar un vector propio $\mathbf{x}_0$ y su valor propio correspondiente λ y hacer

$$\mathbf{x}_k = \lambda^k\mathbf{x}_0 \quad (k = 1, 2, \ldots) \tag{9}$$

Esta sucesión funciona, porque

$$A\mathbf{x}_k = A(\lambda^k\mathbf{x}_0) = \lambda^k(A\mathbf{x}_0) = \lambda^k(\lambda\mathbf{x}_0) = \lambda^{k+1}\mathbf{x}_0 = \mathbf{x}_{k+1}$$

¡Las combinaciones lineales de soluciones de la forma (9) también son soluciones! Vea el ejercicio 33.

PROBLEMAS DE PRÁCTICA

1. ¿5 es un valor propio de $A = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}$?

2. Si $\mathbf{x}$ es un vector propio para A correspondiente a λ, ¿qué es $A^3\mathbf{x}$?

5.1 EJERCICIOS

1. ¿$\lambda = 2$ es un valor propio de $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$? ¿Por qué sí o por qué no?

2. ¿$\lambda = -2$ es un valor propio de $\begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$? ¿Por qué sí o por qué no?

3. ¿$\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ es un vector propio de $\begin{bmatrix} -3 & 1 \\ -3 & 8 \end{bmatrix}$? Si lo es, encuentre el valor propio.

4. ¿$\begin{bmatrix} -1+\sqrt{2} \\ 1 \end{bmatrix}$ es un vector propio de $\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$? Si lo es, encuentre el valor propio.

5. ¿$\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$ es un vector propio de $\begin{bmatrix} 3 & 7 & 9 \\ -4 & -5 & 1 \\ 2 & 4 & 4 \end{bmatrix}$? Si lo es, encuentre el valor propio.

6. ¿$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ es un vector propio de $\begin{bmatrix} 3 & 6 & 7 \\ 3 & 3 & 7 \\ 5 & 6 & 5 \end{bmatrix}$? Si lo es, encuentre el valor propio.

7. ¿$\lambda = 4$ es un valor propio de $\begin{bmatrix} 3 & 0 & -1 \\ 2 & 3 & 1 \\ -3 & 4 & 5 \end{bmatrix}$? Si lo es, encuentre el vector propio correspondiente.

8. ¿$\lambda = 3$ es un valor propio de $\begin{bmatrix} 1 & 2 & 2 \\ 3 & -2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$? Si lo es, encuentre el vector propio correspondiente.

En los ejercicios 9 a 16, encuentre una base para el espacio propio correspondiente a cada valor propio enlistado.

9. $A = \begin{bmatrix} 5 & 0 \\ 2 & 1 \end{bmatrix}, \lambda = 1, 5$

10. $A = \begin{bmatrix} 10 & -9 \\ 4 & -2 \end{bmatrix}, \lambda = 4$

11. $A = \begin{bmatrix} 4 & -2 \\ -3 & 9 \end{bmatrix}, \lambda = 10$

12. $A = \begin{bmatrix} 7 & 4 \\ -3 & -1 \end{bmatrix}, \lambda = 1, 5$

13. $A = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \lambda = 1, 2, 3$

14. $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & -3 & 0 \\ 4 & -13 & 1 \end{bmatrix}, \lambda = -2$

15. $A = \begin{bmatrix} 4 & 2 & 3 \\ -1 & 1 & -3 \\ 2 & 4 & 9 \end{bmatrix}, \lambda = 3$

16. $A = \begin{bmatrix} 3 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \lambda = 4$

Encuentre los valores propios de las matrices dadas en los ejercicios 17 y 18.

17. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 5 \\ 0 & 0 & -1 \end{bmatrix}$

18. $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -3 \end{bmatrix}$

19. Encuentre un valor propio para $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, sin hacer cálculos. Justifique su respuesta.

20. Sin hacer cálculos, encuentre un valor propio y dos vectores propios linealmente independientes de $A = \begin{bmatrix} 5 & 5 & 5 \\ 5 & 5 & 5 \\ 5 & 5 & 5 \end{bmatrix}$. Justifique su respuesta.

En los ejercicios 21 y 22, A es una matriz de $n \times n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

21. a. Si $A\mathbf{x} = \lambda\mathbf{x}$ para algún vector $\mathbf{x}$, entonces λ es un valor propio de A.

b. Una matriz A es no invertible si, y sólo si, 0 es un valor propio de A.

c. Un número c es un valor propio de A si, y sólo si, la ecuación $(A - cI)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial.

d. Puede ser difícil encontrar un vector propio de A, pero es fácil comprobar si un vector dado es un vector propio.

e. Para encontrar los valores propios de A, reduzca A a su forma escalonada.

22. a. Si $A\mathbf{x} = \lambda\mathbf{x}$ para algún escalar λ, entonces $\mathbf{x}$ es un vector propio de A.

b. Si $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios linealmente independientes, entonces corresponden a diferentes valores propios.

c. Un vector de estado estacionario para una matriz estocástica es en realidad un vector propio.

d. Los valores propios de una matriz están en su diagonal principal.

e. Un espacio propio de A es un espacio nulo de cierta matriz.

23. Explique por qué una matriz de 2×2 puede tener, cuando mucho, dos valores propios distintos. Explique por qué una matriz de $n \times n$ puede tener, cuando mucho, n valores propios distintos.

24. Estructure un ejemplo de una matriz de 2×2 que sólo tenga un valor propio distinto.

25. Sea λ un valor propio de una matriz invertible A. Muestre que λ^{-1} es un valor propio de A^{-1}. [*Sugerencia:* Suponga que una $\mathbf{x}$ diferente de cero satisface $A\mathbf{x} = \lambda\mathbf{x}$.]

26. Muestre que si A^2 es la matriz cero, entonces el único valor propio de A es 0.

27. Muestre que λ es un valor propio de A si, y sólo si, λ es un valor propio de A^T. [*Sugerencia:* Encuentre la relación entre $A - \lambda I$ y $A^T - \lambda I$.]

28. Use el ejercicio 27 y complete la demostración del teorema 1 para el caso en que A es triangular inferior.

29. Considere una matriz A de $n \times n$ con la propiedad de que todas las sumas de fila son iguales al mismo número s. Muestre que s es un valor propio de A. [*Sugerencia:* Encuentre un vector propio.]

30. Considere una matriz A de $n \times n$ con la propiedad de que todas las sumas de columna son iguales al mismo número s. Muestre que s es un valor propio de A. [*Sugerencia:* Use los ejercicios 27 y 29.]

En los ejercicios 31 y 32, sea A la matriz de la transformación lineal T. Sin escribir A, encuentre un valor propio de A y describa el espacio propio.

31. T es la transformación en $\mathbb{R}^2$ que refleja puntos sobre alguna línea que pasa por el origen.

32. T es la transformación en $\mathbb{R}^3$ que gira puntos sobre alguna línea que pasa por el origen.

33. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores propios de una matriz A, con valores propios correspondientes λ y μ, y sean c_1 y c_2 escalares. Defina

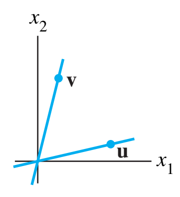
$$\mathbf{x}_k = c_1\lambda^k\mathbf{u} + c_2\mu^k\mathbf{v} \quad (k = 0, 1, 2, \ldots)$$

a. ¿Por definición, qué es $\mathbf{x}_{k+1}$?

b. Calcule $A\mathbf{x}_k$ a partir de la fórmula para $\mathbf{x}_k$ y muestre que $A\mathbf{x}_k = \mathbf{x}_{k+1}$. Este cálculo demostrará que la sucesión $\{\mathbf{x}_k\}$ definida anteriormente satisface la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$.

34. Describa cómo podría intentar construir una solución de una ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$ si estuviera dada la $\mathbf{x}_0$ inicial y este vector resultara no ser un vector propio de A. [*Pista:* ¿Cómo podría relacionarse $\mathbf{x}_0$ con los vectores propios de A?]

35. Sean $\mathbf{u}$ y $\mathbf{v}$ los vectores mostrados en la figura, y suponga que son vectores propios de una matriz A de 2×2 que corresponden a los valores propios 2 y 3, respectivamente. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal dada por $T(\mathbf{x}) = A\mathbf{x}$ para cada $\mathbf{x}$ en $\mathbb{R}^2$, y sea $\mathbf{w} = \mathbf{u} + \mathbf{v}$. Haga una copia de la figura y, sobre el mismo sistema de coordenadas, grafique cuidadosamente los vectores $T(\mathbf{u})$, $T(\mathbf{v})$ y $T(\mathbf{w})$.



36. Repita el ejercicio 35, suponiendo que $\mathbf{u}$ y $\mathbf{v}$ son vectores propios de A que corresponden a los valores propios -1 y 3, respectivamente.

[M] En los ejercicios 37 a 40 use un programa para encontrar los valores propios de la matriz. Después use el método del ejemplo 4 con una rutina de reducción por filas para producir una base para cada espacio propio.

37. $\begin{bmatrix} 8 & -10 & -5 \\ 2 & 17 & 2 \\ -9 & -18 & 4 \end{bmatrix}$

38. $\begin{bmatrix} 9 & -4 & -2 & -4 \\ -56 & 32 & -28 & 44 \\ -14 & -14 & 6 & -14 \\ 42 & -33 & 21 & -45 \end{bmatrix}$

39. $\begin{bmatrix} 4 & -9 & -7 & 8 & 2 \\ -7 & -9 & 0 & 7 & 14 \\ 5 & 10 & 5 & -5 & -10 \\ -2 & 3 & 7 & 0 & 4 \\ -3 & -13 & -7 & 10 & 11 \end{bmatrix}$

40. $\begin{bmatrix} -4 & -4 & 20 & -8 & -1 \\ 14 & 12 & 46 & 18 & 2 \\ 6 & 4 & -18 & 8 & 1 \\ 11 & 7 & -37 & 17 & 2 \\ 18 & 12 & -60 & 24 & 5 \end{bmatrix}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. El número 5 es un valor propio de A si, y sólo si, la ecuación $(A - 5I)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial. Forme

$$A - 5I = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix} - \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 1 \\ 3 & -5 & 5 \\ 2 & 2 & 1 \end{bmatrix}$$

y reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 1 & -3 & 1 & 0 \\ 3 & -5 & 5 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 8 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix}$$

En este punto, es evidente que el sistema homogéneo no tiene variables libres. Entonces $A - 5I$ es una matriz invertible, lo cual significa que 5 *no* es un valor propio de A.

2. Si $\mathbf{x}$ es un vector propio de A correspondiente a λ, entonces $A\mathbf{x} = \lambda\mathbf{x}$, y así sucesivamente

$$A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$$

Una vez más, $A^3\mathbf{x} = A(A^2\mathbf{x}) = A(\lambda^2\mathbf{x}) = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}$. El patrón general, $A^k\mathbf{x} = \lambda^k$, se demuestra por inducción.

5.2 LA ECUACIÓN CARACTERÍSTICA

Información útil acerca de los valores propios de una matriz cuadrada A se encuentra codificada en una ecuación escalar llamada ecuación característica de A. Un ejemplo sencillo conduce al caso general.

EJEMPLO 1 Encuentre los valores propios de $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.

Solución Deben encontrarse todos los escalares λ tales que la ecuación matricial

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

tenga una solución no trivial. De acuerdo con el teorema de la matriz invertible de la sección 2.3, este problema es equivalente a encontrar todas las λ tales que la matriz $A - \lambda I$ *no* sea invertible, donde

$$A - \lambda I = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}$$

Según el teorema 4 de la sección 2.2, esta matriz no es invertible precisamente cuando su determinante es cero. Así que los valores propios de A son las soluciones para la ecuación

$$\det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = 0$$

Recuerde que

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

Así que

$$\begin{aligned} \det(A - \lambda I) &= (2-\lambda)(-6-\lambda) - (3)(3) \\ &= -12 + 6\lambda - 2\lambda + \lambda^2 - 9 \\ &= \lambda^2 + 4\lambda - 21 \end{aligned}$$

Al establecer $\lambda^2 + 4\lambda - 21 = 0$, se tiene que $(\lambda - 3)(\lambda + 7) = 0$; así que los valores propios de A son 3 y -7.

El determinante del ejemplo 1 transformó la ecuación de matrices $(A - \lambda I)\mathbf{x} = \mathbf{0}$, donde intervienen *dos* incógnitas (λ y $\mathbf{x}$), en la ecuación escalar $\lambda^2 + 4\lambda - 21 = 0$, que tiene solamente una incógnita. La misma idea funciona para matrices de $n \times n$. Sin embargo, antes de pasar a matrices mayores, se presenta un resumen de las propiedades de los determinantes necesarios para estudiar los valores propios.

Determinantes

Sean A una matriz de $n \times n$, U cualquier forma escalonada obtenida a partir de A mediante reemplazo e intercambio de filas (sin cambiar de escala), y r la cantidad de tales intercambios de filas. Entonces el **determinante** de A, escrito det A, es $(-1)^r$ veces el producto de las entradas diagonales $u_{11}, \ldots, u_{nn}$ de U. Si A es invertible, entonces $u_{11}, \ldots, u_{nn}$ son todas *pivotes* (porque $A \sim I_n$ y las u_{ii} no se han escalado a números 1). En caso contrario, por lo menos u_{nn} es cero, y el producto $u_{11} \cdots u_{nn}$ es cero. Así que[1]

$$\det A = \begin{cases} (-1)^r \cdot \begin{pmatrix} \text{producto de los} \\ \text{pivotes en } U \end{pmatrix}, & \text{cuando } A \text{ es invertible} \\ 0, & \text{cuando } A \text{ es no invertible} \end{cases} \qquad (1)$$

[1]La fórmula (1) se dedujo en la sección 3.2. Los lectores que no hayan estudiado el capítulo 3 pueden usar esta fórmula como la definición de det A. Es un hecho notable y no trivial que cualquier forma escalonada U obtenida a partir de A sin cambiar la escala proporciona el mismo valor de det A.

EJEMPLO 2 Calcule det A para $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.

Solución La siguiente reducción por filas utiliza un intercambio de filas:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix} = U_1$$

Entonces det A es igual a $(-1)^1(1)(-2)(-1) = -2$. La siguiente reducción por filas alternativa evita el intercambio entre filas y produce una forma escalonada diferente. En el último paso se suma $-1/3$ veces la fila 2 a la fila 3:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & 0 & 1/3 \end{bmatrix} = U_2$$

Esta vez det A es $(-1)^0(1)(-6)(1/3) = -2$, igual que antes.

La fórmula (1) para el determinante muestra que A es invertible si, y sólo si, det A es diferente de cero. Este hecho, y la caracterización de invertibilidad que se encontró en la sección 5.1, pueden agregarse al teorema de la matriz invertible.

TEOREMA

Teorema de la matriz invertible (continuación)

Sea A una matriz de $n \times n$. Entonces A es invertible si, y sólo si:

s. El número 0 *no* es un valor propio de A.

t. El determinante de A *no* es cero.

Si A es una matriz de 3×3, entonces $|\det A|$ resulta ser el *volumen* del paralelepípedo determinado por las columnas $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ de A, como en la figura 1. (Para mayores detalles vea la sección 3.3.) Este volumen es *distinto de cero* si, y sólo si, los vectores $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ son linealmente independientes y la matriz A es invertible. (Si los vectores son distintos de cero y linealmente independientes, pertenecen a un plano o se encuentran a lo largo de una línea.)

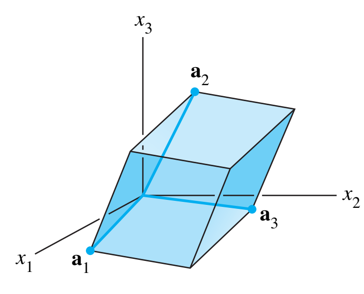


FIGURA 1

El teorema siguiente enlista los datos necesarios de las secciones 3.1 y 3.2. Se incluye el inciso (a) para usarlo como una referencia conveniente.

TEOREMA 3 **Propiedades de los determinantes**

Sean A y B matrices de $n \times n$.

a. A es invertible si, y sólo si, $\det A \neq 0$.
b. $\det AB = (\det A)(\det B)$.
c. $\det A^T = \det A$.
d. Si A es triangular, entonces $\det A$ es el producto de las entradas que están en la diagonal principal de A.
e. Una operación de reemplazo de fila en A no cambia el determinante. Un intercambio de fila cambia el signo del determinante. Un escalamiento de fila también escala el determinante por el mismo factor escalar.

La ecuación característica

En virtud del teorema 3(a), puede utilizarse un determinante para decidir cuándo una matriz $A - \lambda I$ *no* es invertible. La ecuación escalar $\det(A - \lambda I) = 0$ es la **ecuación característica** de A, y el argumento del ejemplo 1 justifica el enunciado siguiente.

Un escalar λ es un valor propio de una matriz A de $n \times n$ si, y sólo si, λ satisface la ecuación característica

$$\det(A - \lambda I) = 0$$

EJEMPLO 3 Encuentre la ecuación característica de

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Solución Forme $A - \lambda I$, y use el teorema 3(d):

$$\det(A - \lambda I) = \det \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix}$$
$$= (5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda)$$

La ecuación característica es

$$(5-\lambda)^2(3-\lambda)(1-\lambda) = 0$$

o bien

$$(\lambda - 5)^2(\lambda - 3)(\lambda - 1) = 0$$

Al desarrollar el producto, también puede escribirse

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0$$

En los ejemplos 1 y 3, det $(A - \lambda I)$ es un polinomio en λ. Puede mostrarse que si A es una matriz de $n \times n$, entonces det $(A - \lambda I)$ es un polinomio de grado n llamado **polinomio característico** de A.

Se afirma que el valor propio 5 del ejemplo 3 tiene *multiplicidad* 2 porque $(\lambda - 5)$ aparece dos veces como factor del polinomio característico. En general, la **multiplicidad** (**algebraica**) de un valor propio λ es su multiplicidad como una raíz de la ecuación característica.

EJEMPLO 4 El polinomio característico de una matriz de 6×6 es $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Encuentre los valores propios y su multiplicidad.

Solución Factorice el polinomio

$$\lambda^6 - 4\lambda^5 - 12\lambda^4 = \lambda^4(\lambda^2 - 4\lambda - 12) = \lambda^4(\lambda - 6)(\lambda + 2)$$

Los valores propios son 0 (multiplicidad 4), 6 (multiplicidad 1), y -2 (multiplicidad 1).

También podrían enlistarse los valores propios del ejemplo 4 como 0, 0, 0, 0, 6 y -2, de manera que los valores propios se repitan de acuerdo a su multiplicidad.

Debido a que la ecuación característica de una matriz de $n \times n$ implica un polinomio de grado n, la ecuación tiene exactamente n raíces, contando las multiplicidades, dado que se permiten raíces complejas. Estas raíces complejas, llamadas *valores propios complejos*, se analizarán en la sección 5.5. Hasta entonces, se considerarán sólo valores propios reales y los escalares seguirán siendo números reales.

La ecuación característica es importante para los propósitos teóricos. En el trabajo práctico, sin embargo, los valores propios de cualquier matriz mayor de 2×2 deben encontrarse por medio de una computadora, a menos que la matriz sea triangular o tenga otras propiedades especiales. Aunque es fácil calcular a mano un polinomio característico de 3×3, puede resultar difícil factorizarlo (a menos que se elija cuidadosamente la matriz). Vea las notas numéricas incluidas al final de esta sección.

SG Factorización de un polinomio 5 a 8 (Factoring a Polynomial 5-8)

Semejanza

El teorema siguiente ilustra un uso del polinomio característico y proporciona la base para varios métodos iterativos que *aproximan* valores propios. Si A y B son matrices de $n \times n$, entonces A **es semejante a** B si existe una matriz invertible P tal que $P^{-1}AP = B$, o, de manera equivalente, $A = PBP^{-1}$. Si se escribe Q en vez de P^{-1}, se tiene $Q^{-1}BQ = A$. Así que B también es semejante a A, y simplemente se dice que A y B **son semejantes**. Cuando A se convierte en $P^{-1}AP$ se realiza una **transformación de semejanza**.

TEOREMA 4 Si las matrices A y B de $n \times n$ son semejantes, entonces tienen el mismo polinomio característico y, por lo tanto, los mismos valores propios (con las mismas multiplicidades).

DEMOSTRACIÓN Si $B = P^{-1}AP$, entonces

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

Con el uso de la propiedad multiplicativa (b) del teorema 3, se calcula

$$\begin{aligned} \det(B - \lambda I) &= \det[P^{-1}(A - \lambda I)P] \\ &= \det(P^{-1}) \cdot \det(A - \lambda I) \cdot \det(P) \end{aligned} \tag{2}$$

Como $\det(P^{-1}) \cdot \det(P) = \det(P^{-1}P) = \det I = 1$, se observa en (2) que $\det(B - \lambda I) = \det(A - \lambda I)$. ■

Advertencia: Semejanza no es lo mismo que equivalencia por filas. (Si A es equivalente por filas a B, entonces $B = EA$ para alguna matriz invertible E.) Las operaciones por fila sobre una matriz normalmente alteran sus valores propios.

Aplicación a los sistemas dinámicos

Los valores propios y los vectores propios tienen la clave para la evolución discreta de un sistema dinámico, tal como se mencionó en la introducción del capítulo.

EJEMPLO 5 Sea $A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. Analice el comportamiento a largo plazo del sistema dinámico definido por $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0, 1, 2, \ldots$), con $\mathbf{x}_0 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$.

Solución El primer paso es encontrar los valores propios de A y una base para cada espacio propio. La ecuación característica de A es

$$\begin{aligned} 0 &= \det\begin{bmatrix} .95 - \lambda & .03 \\ .05 & .97 - \lambda \end{bmatrix} = (.95 - \lambda)(.97 - \lambda) - (.03)(.05) \\ &= \lambda^2 - 1.92\lambda + .92 \end{aligned}$$

Por la fórmula cuadrática

$$\begin{aligned} \lambda &= \frac{1.92 \pm \sqrt{(1.92)^2 - 4(.92)}}{2} = \frac{1.92 \pm \sqrt{.0064}}{2} \\ &= \frac{1.92 \pm .08}{2} = 1 \text{ o bien } .92 \end{aligned}$$

Se comprueba rápidamente que los vectores propios correspondientes a $\lambda = 1$ y $\lambda = .92$ son múltiplos de

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

respectivamente.

El siguiente paso es escribir la $\mathbf{x}_0$ dada en términos de $\mathbf{v}_1$ y $\mathbf{v}_2$. Esto puede hacerse porque $\{\mathbf{v}_1, \mathbf{v}_2\}$ es, evidentemente, una base para $\mathbb{R}^2$. (¿Por qué?) Así, existen pesos c_1 y c_2 tales que

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = [\,\mathbf{v}_1 \quad \mathbf{v}_2\,] \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{3}$$

De hecho,

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= [\,\mathbf{v}_1 \quad \mathbf{v}_2\,]^{-1}\mathbf{x}_0 = \begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}^{-1} \begin{bmatrix} .60 \\ .40 \end{bmatrix} \\ &= \frac{1}{-8}\begin{bmatrix} -1 & -1 \\ -5 & 3 \end{bmatrix}\begin{bmatrix} .60 \\ .40 \end{bmatrix} = \begin{bmatrix} .125 \\ .225 \end{bmatrix} \end{aligned} \tag{4}$$

Como $\mathbf{v}_1$ y $\mathbf{v}_2$ en (3) son vectores propios de A, con $A\mathbf{v}_1 = \mathbf{v}_1$ y $A\mathbf{v}_2 = .92\mathbf{v}_2$, se calcula fácilmente cada $\mathbf{x}_k$:

$$\begin{aligned} \mathbf{x}_1 = A\mathbf{x}_0 &= c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 && \text{Usando la linealidad de } \mathbf{x} \mapsto A\mathbf{x} \\ &= c_1\mathbf{v}_1 + c_2(.92)\mathbf{v}_2 && \mathbf{v}_1 \text{ y } \mathbf{v}_2 \text{ son vectores propios} \\ \mathbf{x}_2 = A\mathbf{x}_1 &= c_1A\mathbf{v}_1 + c_2(.92)A\mathbf{v}_2 \\ &= c_1\mathbf{v}_1 + c_2(.92)^2\mathbf{v}_2 \end{aligned}$$

y así sucesivamente. En general,

$$\mathbf{x}_k = c_1\mathbf{v}_1 + c_2(.92)^k\mathbf{v}_2 \quad (k = 0, 1, 2, \ldots)$$

Al usar c_1 y c_2 de (4),

$$\mathbf{x}_k = .125\begin{bmatrix} 3 \\ 5 \end{bmatrix} + .225(.92)^k\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (k = 0, 1, 2, \ldots) \tag{5}$$

Esta fórmula explícita para $\mathbf{x}_k$ proporciona la solución de la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Cuando $k \to \infty$, $(.92)^k$ tiende a cero y $\mathbf{x}_k$ tiende a $\begin{bmatrix} .375 \\ .625 \end{bmatrix} = .125\mathbf{v}_1$.

Los cálculos del ejemplo 5 tienen una aplicación interesante a una cadena de Markov en la sección 4.9. Quien haya leído esa sección tal vez reconozca que A en el ejemplo 5 anterior es la misma que la matriz de migración M vista en la sección 4.9, $\mathbf{x}_0$ es la distribución de población inicial entre la ciudad y los suburbios, y $\mathbf{x}_k$ representa la distribución de población después de k años.

El teorema 18 de la sección 4.9 afirma que para una matriz como A, la sucesión $\mathbf{x}_k$ tiende a un vector de estado estacionario. Ahora se sabe *por qué* $\mathbf{x}_k$ se comporta de esta manera, al menos para la matriz de migración. El vector de estado estacionario es $.125\mathbf{v}_1$, un múltiplo del vector propio $\mathbf{v}_1$, y la fórmula (5) para $\mathbf{x}_k$ muestra exactamente por qué $\mathbf{x}_k \to .125\mathbf{v}_1$.

NOTAS NUMÉRICAS

1. Programas de computadora como Mathematica y Maple puede usar cálculos simbólicos para encontrar el polinomio característico de una matriz de tamaño moderado. Pero no hay una fórmula o algoritmo finito para resolver la ecuación característica de una matriz general de $n \times n$ para $n \geq 5$.
2. Los mejores métodos numéricos para encontrar los valores propios evitan por completo a los polinomios característicos. De hecho, MATLAB encuentra el polinomio característico de una matriz A calculando primero los valores propios $\lambda_1, \ldots, \lambda_n$ de A, y desarrollando luego el producto $(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$.
3. Varios algoritmos comunes para calcular los valores propios de una matriz A se basan en el teorema 4. El poderoso *algoritmo QR* se analiza en los ejercicios. Otra técnica, llamada *método de Jacobi*, funciona cuando $A = A^T$ y calcula una sucesión de matrices de la forma

$$A_1 = A \quad \text{y} \quad A_{k+1} = P_k^{-1} A_k P_k \quad (k = 1, 2, \ldots)$$

   Cada matriz de la sucesión es semejante a A y, por lo tanto, tiene los mismos valores propios que A. Las entradas no diagonales de A_{k+1} tienden a cero al aumentar k, y las entradas diagonales tienden a aproximar los valores propios de A.
4. En la sección 5.8 se analizan otros métodos para calcular valores propios.

PROBLEMA DE PRÁCTICA

Encuentre la ecuación característica y los valores propios de $A = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$.

5.2 EJERCICIOS

En los ejercicios 1 a 8, encuentre el polinomio característico y los valores propios de las matrices dadas.

1. $\begin{bmatrix} 2 & 7 \\ 7 & 2 \end{bmatrix}$ **2.** $\begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$

3. $\begin{bmatrix} 3 & -2 \\ 1 & -1 \end{bmatrix}$ **4.** $\begin{bmatrix} 5 & -3 \\ -4 & 3 \end{bmatrix}$

5. $\begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$ **6.** $\begin{bmatrix} 3 & -4 \\ 4 & 8 \end{bmatrix}$

7. $\begin{bmatrix} 5 & 3 \\ -4 & 4 \end{bmatrix}$ **8.** $\begin{bmatrix} 7 & -2 \\ 2 & 3 \end{bmatrix}$

Los ejercicios 9 a 14 requieren técnicas de la sección 3.1. Encuentre el polinomio característico de cada matriz, usando ya sea un desarrollo por cofactores o la fórmula especial para los determinantes 3×3 descrita antes de los ejercicios 15 a 18 en la sección 3.1. [*Nota:* No es fácil encontrar el polinomio característico de una matriz de 3×3 sólo con operaciones por fila, porque interviene la variable λ.]

9. $\begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & -1 \\ 0 & 6 & 0 \end{bmatrix}$ **10.** $\begin{bmatrix} 0 & 3 & 1 \\ 3 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}$

11. $\begin{bmatrix} 4 & 0 & 0 \\ 5 & 3 & 2 \\ -2 & 0 & 2 \end{bmatrix}$ **12.** $\begin{bmatrix} -1 & 0 & 1 \\ -3 & 4 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

13. $\begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 5 & -2 & 3 \\ 0 & 1 & 0 \\ 6 & 7 & -2 \end{bmatrix}$

Para las matrices de los ejercicios 15 a 17 enliste los valores propios, repetidos de acuerdo con sus multiplicidades.

15. $\begin{bmatrix} 4 & -7 & 0 & 2 \\ 0 & 3 & -4 & 6 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ **16.** $\begin{bmatrix} 5 & 0 & 0 & 0 \\ 8 & -4 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 1 & -5 & 2 & 1 \end{bmatrix}$

17. $\begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 & 0 \\ 3 & 8 & 0 & 0 & 0 \\ 0 & -7 & 2 & 1 & 0 \\ -4 & 1 & 9 & -2 & 3 \end{bmatrix}$

18. Puede mostrarse que la multiplicidad algebraica de un valor propio λ siempre es mayor que, o igual a, la dimensión del espacio propio correspondiente a λ. Encuentre h en la matriz A siguiente de manera que el espacio propio para $\lambda = 5$ sea bidimensional:

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & h & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

19. Sea A una matriz de $n \times n$ y suponga que A tiene n valores propios reales, $\lambda_1, \ldots, \lambda_n$, repetidos de acuerdo con sus multiplicidades, de manera que

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$$

Explique por qué $\det A$ es el producto de los n valores propios de A. (Este resultado es válido para cualquier matriz cuadrada cuando se consideren valores propios complejos.)

20. Utilice una propiedad de los determinantes para mostrar que A y A^T tienen el mismo polinomio característico.

En los ejercicios 21 y 22, A y B son matrices de $n \times n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

21. a. El determinante de A es el producto de las entradas diagonales de A.

b. Una operación elemental por filas con A no cambia el determinante.

c. $(\det A)(\det B) = \det AB$.

d. Si $\lambda + 5$ es un factor del polinomio característico de A, entonces 5 es un valor propio de A.

22. a. Si A es de 3×3, con columnas $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$, entonces $\det A$ es igual al volumen del paralelepípedo determinado por $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$.

b. $\det A^T = (-1)\det A$.

c. La multiplicidad de una raíz r de la ecuación característica de A es la multiplicidad algebraica de r como valor propio de A.

d. Una operación de reemplazo de filas con A no cambia los valores propios.

Un método ampliamente utilizado para estimar los valores propios de una matriz A general es el *algoritmo QR*. En condiciones adecuadas, este algoritmo produce una sucesión de matrices, todas semejantes a A, que se vuelven casi triangulares superiores, con entradas diagonales que se aproximan a los valores propios de A. La idea principal es factorizar A (u otra matriz semejante) en la forma $A = Q_1R_1$, donde $Q_1^T = Q_1^{-1}$ y R_1 es triangular superior. Los factores se intercambian para formar $A_1 = R_1Q_1$, que de nuevo se factoriza como $A_1 = Q_2R_2$; luego para formar $A_2 = R_2Q_2$, y así sucesivamente. La semejanza de $A, A_1, \ldots$ se deduce a partir del resultado más general del ejercicio 23.

23. Muestre que si $A = QR$ con Q invertible, entonces A es semejante a $A_1 = RQ$.

24. Muestre que si A y B son semejantes, entonces $\det A = \det B$.

25. Sean $A = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$, $\mathbf{x}_0 = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$. [*Nota:* A es la matriz estocástica estudiada en el ejemplo 5 de la sección 4.9.]

a. Encuentre una base para $\mathbb{R}^2$ que consista en $\mathbf{v}_1$ y otro vector propio $\mathbf{v}_2$ de A.

b. Demuestre que $\mathbf{x}_0$ puede escribirse como una combinación lineal de la forma $\mathbf{x}_0 = \mathbf{v}_1 + c\mathbf{v}_2$.

c. Para $k = 1, 2, \ldots$, defina $\mathbf{x}_k = A^k\mathbf{x}_0$. Calcule $\mathbf{x}_1$ y $\mathbf{x}_2$, y escriba una fórmula para $\mathbf{x}_k$. Luego muestre $\mathbf{x}_k \to \mathbf{v}_1$ conforme k aumenta.

26. Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Use la fórmula (1) para un determinante (dada antes del ejemplo 2) y muestre que $\det A = ad - bc$. Considere dos casos: $a \neq 0$ y $a = 0$.

27. Sean $A = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, y $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

a. Muestre que $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ son vectores propios de A. [*Nota:* A es la matriz estocástica estudiada en el ejemplo 3 de la sección 4.9.]

b. Sea $\mathbf{x}_0$ cualquier vector en $\mathbb{R}^3$ con entradas no negativas cuya suma es 1. (En la sección 4.9, se llamó a $\mathbf{x}_0$ un vector de probabilidad.) Explique por qué hay constantes c_1, c_2, c_3 tales que $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Calcule $\mathbf{w}^T\mathbf{x}_0$, y deduzca que $c_1 = 1$.

c. Para $k = 1, 2, \ldots$, defina $\mathbf{x}_k = A^k\mathbf{x}_0$, con $\mathbf{x}_0$ como en el inciso (b). Muestre que $\mathbf{x}_k \to \mathbf{v}_1$ al aumentar k.

28. **[M]** Construya una matriz aleatoria A de 4×4 con valores enteros, y verifique si A y A^T tienen el mismo polinomio característico (los mismos valores propios con la misma multiplicidad). ¿Tienen A y A^T los mismos vectores propios? Efectúe el mismo análisis con una matriz de 5×5. Escriba las matrices e informe acerca de sus conclusiones.

29. **[M]** Construya una matriz aleatoria A de 4×4 con valores enteros.

a. Reduzca A a la forma escalonada U sin escalar las filas, y use U en la fórmula (1) para calcular det A. (Si A resulta ser singular, comience de nuevo con otra matriz aleatoria.)

b. Determine los valores propios de A y el producto de estos valores propios (con la mayor precisión posible).

c. Enliste la matriz A y, con cuatro cifras decimales, enumere los pivotes de U y los valores propios de A. Calcule det A con un programa de matrices y compárelo con los productos encontrados en (a) y (b).

30. **[M]** Sea $A = \begin{bmatrix} -6 & 28 & 21 \\ 4 & -15 & -12 \\ -8 & a & 25 \end{bmatrix}$. Para cada valor de a en el conjunto $\{32, 31.9, 31.8, 32.1, 32.2\}$, calcule el polinomio característico de A y los valores propios. En cada caso, trace una gráfica del polinomio característico $p(t) = \det(A - tI)$ para $0 \le t \le 3$. De ser posible, estructure todas las gráficas en un solo sistema de coordenadas. Describa la forma en que las gráficas revelan los cambios de los valores propios al cambiar a.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

La ecuación característica es

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & -4 \\ 4 & 2-\lambda \end{bmatrix}$$
$$= (1-\lambda)(2-\lambda) - (-4)(4) = \lambda^2 - 3\lambda + 18$$

A partir de la fórmula cuadrática,

$$\lambda = \frac{3 \pm \sqrt{(-3)^2 - 4(18)}}{2} = \frac{3 \pm \sqrt{-63}}{2}$$

Es claro que la ecuación característica no tiene soluciones reales, de manera que A no tiene valores propios reales. La matriz A está actuando sobre el espacio vectorial real $\mathbb{R}^2$, y no existe un vector $\mathbf{v}$ diferente de cero en $\mathbb{R}^2$ tal que $A\mathbf{v} = \lambda\mathbf{v}$ para algún escalar λ.

5.3 DIAGONALIZACIÓN

En muchos casos, la información vector propio-valor propio contenida dentro de una matriz A puede representarse mediante una útil factorización de la forma $A = PDP^{-1}$. En esta sección, la factorización permite calcular rápidamente A^k para valores grandes de k, una idea fundamental en varias aplicaciones del álgebra lineal. Posteriormente, en las secciones 5.6 y 5.7, la factorización se usará para analizar (y *desacoplar*) sistemas dinámicos.

En la factorización, la D significa *diagonal*. El cálculo de las potencias de una D de este tipo es trivial.

EJEMPLO 1 Si $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, entonces $D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$ y

$$D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

En general,

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \quad \text{para } k \ge 1$$

Si $A = PDP^{-1}$ para alguna P invertible y una diagonal D, entonces A^k también es fácil de calcular, como se muestra en el siguiente ejemplo.

EJEMPLO 2 Sea $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Encuentre una fórmula para A^k, dado que $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

Solución De la fórmula estándar para el inverso de una matriz de 2×2 se obtiene

$$P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Entonces, por la asociatividad de la multiplicación de matrices,

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD\underbrace{(P^{-1}P)}_{I}DP^{-1} = PDDP^{-1}$$

$$= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

De nuevo,

$$A^3 = (PDP^{-1})A^2 = (PD\underbrace{P^{-1})P}_{I}D^2P^{-1} = PDD^2P^{-1} = PD^3P^{-1}$$

En general, para $k \geq 1$,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}$$

Se dice que una matriz cuadrada A es **diagonalizable** si A es semejante a una matriz diagonal, esto es, si $A = PDP^{-1}$ para alguna matriz P invertible y alguna matriz diagonal D. El siguiente teorema proporciona una caracterización de las matrices diagonalizables e indica la forma de estructurar una factorización adecuada.

TEOREMA 5 **El teorema de la diagonalización**

Una matriz A de $n \times n$ es diagonalizable si, y sólo si, A tiene n vectores propios linealmente independientes.

De hecho, $A = PDP^{-1}$, con D como una matriz diagonal, si, y sólo si, las columnas de P son n vectores propios de A linealmente independientes. En este caso, las entradas diagonales de D son valores propios de A que corresponden, respectivamente, a los vectores propios de P.

En otras palabras, A es diagonalizable si, y sólo si, hay suficientes vectores propios para formar una base de $\mathbb{R}^n$. A una base de este tipo se le denomina **base de vectores propios**.

DEMOSTRACIÓN Primero, observe que si P es cualquier matriz de $n \times n$ con columnas $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y si D es cualquier matriz diagonal con entradas diagonales $\lambda_1, \ldots, \lambda_n$, entonces

$$AP = A[\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n\,] = [\,A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n\,] \tag{1}$$

mientras que

$$PD = P\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\,\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n\,] \tag{2}$$

Ahora suponga que A es diagonalizable e igual a PDP^{-1}. Entonces, al multiplicar por la derecha esta relación por P, se tendrá $AP = PD$. En este caso, (1) y (2) implican que

$$[\,A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n\,] = [\,\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n\,] \tag{3}$$

Igualando columnas, se encuentra que

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \quad A\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \quad \ldots, \quad A\mathbf{v}_n = \lambda_n\mathbf{v}_n \tag{4}$$

Como P es invertible, sus columnas $\mathbf{v}_1, \ldots, \mathbf{v}_n$ deben ser linealmente independientes. También, como estas columnas son diferentes de cero, (4) muestra que $\lambda_1, \ldots, \lambda_n$ son valores propios y que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ son los vectores propios correspondientes. Este argumento demuestra las partes "sólo si" del primero, segundo y tercer enunciados del teorema.

Por último, dados cualesquiera n vectores propios $\mathbf{v}_1, \ldots, \mathbf{v}_n$, úselos para estructurar las columnas de P y utilice los valores propios correspondientes $\lambda_1, \ldots, \lambda_n$ para conformar D. De acuerdo con (1), (2) y (3), $AP = PD$. Esto es válido sin condición alguna para los vectores propios. Si, de hecho, los vectores propios son linealmente independientes, entonces P es invertible (por el teorema de la matriz invertible), y $AP = PD$ implica que $A = PDP^{-1}$. ■

Diagonalización de matrices

EJEMPLO 3 Diagonalice la siguiente matriz, si es posible.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Esto es, encuentre una matriz invertible P y una matriz diagonal D tales que $A = PDP^{-1}$.

Solución Se requieren cuatro pasos para implementar la descripción del teorema 5.

Paso 1. Encontrar los valores propios de A. Como se mencionó en la sección 5.2, las mecánicas a seguir en este paso son apropiadas para una computadora cuando la matriz es mayor de 2×2. Para evitar distracciones inútiles, por lo general, el texto proporcionará la información necesaria para cubrir este paso.

En el presente caso, resulta que la ecuación característica contiene un polinomio cúbico al cual se puede factorizar:

$$\begin{aligned} 0 = \det(A - \lambda I) &= -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

Los valores propios son $\lambda = 1$ y $\lambda = -2$.

Paso 2. Encontrar tres vectores propios de A linealmente independientes. Se necesitan *tres* vectores porque A es una matriz de 3×3. Éste es el paso crítico. Si falla, entonces el teorema 5 postula que A no puede diagonalizarse. El método de la sección 5.1 produce una base para cada espacio propio:

$$\text{Base para } \lambda = 1: \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -2: \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Puede comprobarse que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto linealmente independiente.

Paso 3. Estructurar P a partir de los vectores del paso 2. El orden de los vectores no tiene importancia. Al usar el orden elegido en el paso 2, forma

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Paso 4. Estructurar D a partir de los valores propios correspondientes. En este paso, resulta esencial que el orden de los valores propios corresponda al orden elegido para las columnas de P. Utilice el valor propio $\lambda = -2$ dos veces, una para cada uno de los vectores propios correspondientes a $\lambda = -2$:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Es recomendable comprobar que P y D realmente funcionen. Para evitar calcular P^{-1}, simplemente verifique si $AP = PD$. Esto equivale a $A = PDP^{-1}$ cuando P es invertible. (Sin embargo, ¡compruebe que P sea invertible!) Se calcula

$$AP = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

$$PD = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$ ■

EJEMPLO 4 Diagonalice la siguiente matriz, si es posible.

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Solución La ecuación característica de A resulta ser exactamente la misma que la del ejemplo 3:

$$0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$$

Los valores propios son $\lambda = 1$ y $\lambda = -2$. Sin embargo, al buscar los vectores propios, se encuentra que cada uno de los espacios propios tiene sólo una dimensión:

$$\text{Base para } \lambda = 1: \qquad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -2: \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

No existen otros valores propios, y cada vector propio de A es un múltiplo ya sea de $\mathbf{v}_1$ o de $\mathbf{v}_2$. Por lo tanto, es imposible construir una base de $\mathbb{R}^3$ usando vectores propios de A. De acuerdo con el teorema 5, A *no* es diagonalizable.

El teorema siguiente proporciona una condición *suficiente* para que una matriz sea diagonalizable.

TEOREMA 6 Una matriz de $n \times n$ con n valores propios distintos es diagonalizable.

DEMOSTRACIÓN Sean $\mathbf{v}_1, \ldots, \mathbf{v}_n$ los vectores propios correspondientes a los n valores propios distintos de una matriz A. Entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es linealmente independiente según el teorema 2 de la sección 5.1. Por lo tanto A es diagonalizable, de acuerdo con el teorema 5.

Para ser diagonalizable, no es *necesario* que una matriz de $n \times n$ tenga n valores propios distintos. La matriz de 3×3 del ejemplo 3 es diagonalizable aunque sólo tenga dos valores propios distintos.

EJEMPLO 5 Determine si la siguiente matriz es diagonalizable.

$$A = \begin{bmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{bmatrix}$$

Solución ¡Esto es fácil! Dado que la matriz es triangular, sus valores propios son, evidentemente, 5, 0 y -2. Puesto que A es una matriz de 3×3 con tres valores propios distintos, A es diagonalizable.

Matrices cuyos valores propios no son distintos

Si una matriz A de $n \times n$ tiene n valores propios distintos, con vectores propios correspondientes $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y si $P = [\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_n]$, entonces P es automáticamente invertible porque sus columnas son linealmente independientes, según el teorema 2. Cuando A es diagonalizable, pero tiene menos de n valores propios distintos, aún es posible estructurar a P de alguna forma que la vuelva automáticamente invertible, como lo muestra el teorema siguiente.[1]

TEOREMA 7

Sea A una matriz de $n \times n$ cuyos valores propios distintos son $\lambda_1, \ldots, \lambda p$.

a. Para $1 \leq k \leq p$, la dimensión del espacio propio para λ_k es menor o igual que la multiplicidad del valor propio λ_k.

b. La matriz A es diagonalizable si, y sólo si, la suma de las dimensiones de los distintos espacios propios es igual a n, y esto sucede si, y sólo si, la dimensión del espacio propio para cada λ_k es igual a la multiplicidad de λ_k.

c. Si A es diagonalizable y $\mathcal{B}_k$ es una base para el espacio correspondiente a λ_k para cada k, entonces la colección total de vectores en los conjuntos $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forma una base de vectores propios para $\mathbb{R}^n$.

EJEMPLO 6 Diagonalice la siguiente matriz, si es posible.

$$A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix}$$

Solución Como A es una matriz triangular, los valores propios son 5 y -3, cada uno con multiplicidad 2. Usando el método de la sección 5.1, se encuentra una base para cada espacio propio.

$$\text{Base para } \lambda = 5: \quad \mathbf{v}_1 = \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -3: \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

[1] La demostración del teorema 7 es un tanto larga, pero no difícil. Por ejemplo, vea S. Friedberg, A. Insel, y L. Spence, *Linear Algebra*, 3a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 1997), págs. 234-238.

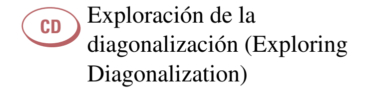
De acuerdo con el teorema 7, el conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es linealmente independiente. Así que la matriz $P = [\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_4]$ es invertible, y $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

PROBLEMAS DE PRÁCTICA

1. Calcule A^8, donde $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$.

2. Sean $A = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, y $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suponga estar enterado de que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios de A. Utilice esta información para diagonalizar A.



Exploración de la diagonalización (Exploring Diagonalization)

3. Sea A una matriz de 4×4 con valores propios 5, 3 y -2, y suponga saber que el espacio propio para $\lambda = 3$ es bidimensional. ¿Tiene usted la suficiente información como para determinar si A es diagonalizable?

WEB

5.3 EJERCICIOS

En los ejercicios 1 y 2, sea $A = PDP^{-1}$ y calcule A^4.

1. $P = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

2. $P = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$

En los ejercicios 3 y 4, utilice la factorización $A = PDP^{-1}$ para calcular A^k, donde k representa un entero positivo arbitrario.

3. $\begin{bmatrix} a & 0 \\ 3(a-b) & b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$

4. $\begin{bmatrix} -2 & 12 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 1 & -3 \end{bmatrix}$

En los ejercicios 5 y 6, la matriz A está factorizada en la forma PDP^{-1}. Use el teorema de la diagonalización para encontrar los valores propios de A y una base para cada espacio propio.

5. $\begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & -3/4 \\ 1/4 & -1/2 & 1/4 \end{bmatrix}$

6. $\begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} -2 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 4 \\ -1 & 0 & -2 \end{bmatrix}$

Diagonalice las matrices de los ejercicios 7 a 20. Los valores propios para los ejercicios 11 a 16 son los siguientes: (11) $\lambda = 1, 2, 3$; (12) $\lambda = 2, 8$; (13) $\lambda = 5, 1$; (14) $\lambda = 5, 4$; (15) $\lambda = 3, 1$; (16) $\lambda = 2, 1$. Para el ejercicio 18, un valor propio es $\lambda = 5$ y un vector propio es $(-2, 1, 2)$.

7. $\begin{bmatrix} 1 & 0 \\ 6 & -1 \end{bmatrix}$

8. $\begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$

9. $\begin{bmatrix} 3 & -1 \\ 1 & 5 \end{bmatrix}$

10. $\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$

11. $\begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}$

12. $\begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$

13. $\begin{bmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & -2 & 2 \end{bmatrix}$

14. $\begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix}$

15. $\begin{bmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{bmatrix}$

16. $\begin{bmatrix} 0 & -4 & -6 \\ -1 & 0 & -3 \\ 1 & 2 & 5 \end{bmatrix}$

17. $\begin{bmatrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

18. $\begin{bmatrix} -7 & -16 & 4 \\ 6 & 13 & -2 \\ 12 & 16 & 1 \end{bmatrix}$

19. $\begin{bmatrix} 5 & -3 & 0 & 9 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

20. $\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$

En los ejercicios 21 y 22, A, B, P y D son matrices de $n \times n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas. (Estudie con cuidado los teoremas 5 y 6 y los ejemplos de esta sección antes de intentar resolver estos ejercicios.)

21. a. A es diagonalizable si $A = PDP^{-1}$ para alguna matriz D y alguna matriz invertible P.

b. Si $\mathbb{R}^n$ tiene una base de vectores propios de A, entonces A es diagonalizable.

c. A es diagonalizable si, y sólo si, tiene n valores propios, contando las multiplicidades.

d. Si A es diagonalizable, entonces es invertible.

22. a. A es diagonalizable si tiene n vectores propios.

b. Si A es diagonalizable, entonces tiene n valores propios distintos.

c. Si $AP = PD$, con D como diagonal, entonces las columnas de P diferentes de cero deben ser vectores propios de A.

d. Si A es invertible, entonces es diagonalizable.

23. A es una matriz de 5×5 y tiene dos valores propios. Un espacio propio es tridimensional y el otro bidimensional. ¿A es diagonalizable? ¿Por qué?

24. A es una matriz de 3×3 con dos valores propios. Cada espacio propio es unidimensional. ¿A es diagonalizable? ¿Por qué?

25. A es una matriz de 4×4 con tres valores propios. Un espacio propio es unidimensional y uno de los otros espacios propios es bidimensional. ¿Es posible que A *no* sea diagonalizable? Justifique su respuesta.

26. A es una matriz de 7×7 con tres valores propios. Un espacio propio es bidimensional y uno de los otros espacios propios es tridimensional. ¿Es posible que A *no* sea diagonalizable? Justifique su respuesta.

27. Muestre que si A es tanto diagonalizable como invertible, entonces también lo es A^{-1}.

28. Muestre que si A tiene n vectores propios linealmente independientes, también los tiene A^T. [*Sugerencia:* Use el teorema de la diagonalización.]

29. Una factorización $A = PDP^{-1}$ no es única. Demuestre esto para la matriz A del ejemplo 2. Con $D_1 = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$, use la información del ejemplo 2 para encontrar una matriz P_1 tal que $A = P_1D_1P_1^{-1}$.

30. Con A y D como en el ejemplo 2, encuentre una P_2 invertible distinta de la P del ejemplo 2, de modo que $A = P_2DP_2^{-1}$.

31. Estructure una matriz de 2×2 distinta de cero que sea invertible pero no diagonalizable.

32. Estructure una matriz de 2×2 no diagonal que sea diagonalizable pero no invertible.

[M] Diagonalice las matrices de los ejercicios 33 a 36. Use el comando de valores propios de un programa de matrices para encontrar los valores propios, y luego determine las bases para los espacios propios como en la sección 5.1.

33. $\begin{bmatrix} -6 & 4 & 0 & 9 \\ -3 & 0 & 1 & 6 \\ -1 & -2 & 1 & 0 \\ -4 & 4 & 0 & 7 \end{bmatrix}$

34. $\begin{bmatrix} 0 & 13 & 8 & 4 \\ 4 & 9 & 8 & 4 \\ 8 & 6 & 12 & 8 \\ 0 & 5 & 0 & -4 \end{bmatrix}$

35. $\begin{bmatrix} 11 & -6 & 4 & -10 & -4 \\ -3 & 5 & -2 & 4 & 1 \\ -8 & 12 & -3 & 12 & 4 \\ 1 & 6 & -2 & 3 & -1 \\ 8 & -18 & 8 & -14 & -1 \end{bmatrix}$

36. $\begin{bmatrix} 4 & 4 & 2 & 3 & -2 \\ 0 & 1 & -2 & -2 & 2 \\ 6 & 12 & 11 & 2 & -4 \\ 9 & 20 & 10 & 10 & -6 \\ 15 & 28 & 14 & 5 & -3 \end{bmatrix}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. $\det(A - \lambda I) = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$. Los valores propios son 2 y 1, y los vectores propios correspondientes son $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Luego, forme

$$P = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \text{y} \qquad P^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

Puesto que $A = PDP^{-1}$,

$$\begin{aligned} A^8 = PD^8P^{-1} &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix} \end{aligned}$$

2. Calcule $A\mathbf{v}_1 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1\cdot\mathbf{v}_1$, y

$$A\mathbf{v}_2 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3\cdot\mathbf{v}_2$$

Por lo tanto, $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios para los valores propios 1 y 3, respectivamente. Entonces

$$A = PDP^{-1}, \quad \text{donde} \quad P = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

3. Sí, A es diagonalizable. Existe una base $\{\mathbf{v}_1, \mathbf{v}_2\}$ para el espacio propio correspondiente a $\lambda = 3$. Además, habrá por lo menos un vector propio para $\lambda = 5$ y uno para $\lambda = -2$. Estos vectores se llamarán $\mathbf{v}_3$ y $\mathbf{v}_4$. Entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es linealmente independiente y, de acuerdo con el teorema 7, A es diagonalizable. No puede haber vectores propios adicionales que sean linealmente independientes de $\mathbf{v}_1, \ldots, \mathbf{v}_4$, porque todos los vectores están en $\mathbb{R}^4$. Así que ambos espacios propios para $\lambda = 5$ y $\lambda = -2$ son unidimensionales.

SG Dominio de los valores propios y los espacios propios 5 a 15 (Mastering: Eigenvalue and Eigenspace 5-15)

5.4 VECTORES PROPIOS Y TRANSFORMACIONES LINEALES

La meta de esta sección es entender la factorización de matrices $A = PDP^{-1}$ como un enunciado acerca de las transformaciones lineales. Se verá que la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es esencialmente la misma que la muy simple función $\mathbf{u} \mapsto D\mathbf{u}$, cuando se observa desde la perspectiva adecuada. Una interpretación de este tipo se aplicará a A y D aun cuando D no es una matriz diagonal.

De la sección 1.9, recuerde que cualquier transformación lineal T de $\mathbb{R}^n$ a $\mathbb{R}^m$ puede implementarse con la multiplicación izquierda por una matriz A, llamada *matriz estándar* de T. Ahora se necesita el mismo tipo de representación para cualquier transformación lineal entre dos espacios vectoriales de dimensión finita.

La matriz de una transformación lineal

Sean V un espacio vectorial n-dimensional, W un espacio vectorial m-dimensional, y T cualquier transformación lineal de V a W. Para asociar una matriz con T, seleccione bases (ordenadas) $\mathcal{B}$ y $\mathcal{C}$ para V y W, respectivamente.

Dado cualquier $\mathbf{x}$ en V, el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ está en $\mathbb{R}^n$ y el vector de coordenadas de su imagen, $[T(\mathbf{x})]_{\mathcal{C}}$, está en $\mathbb{R}^m$, como indica la figura 1.

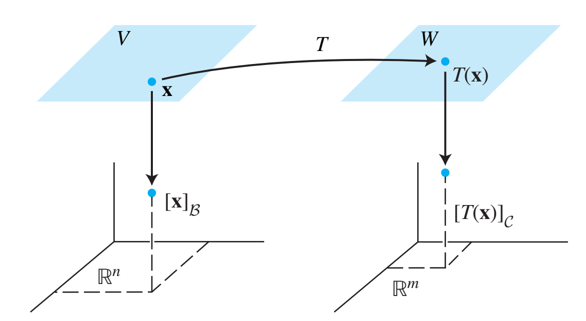


FIGURA 1 Una transformación lineal de V a W.

La conexión entre $[\mathbf{x}]_{\mathcal{B}}$ y $[T(\mathbf{x})]_{\mathcal{C}}$ es fácil de encontrar. Sea $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ la base $\mathcal{B}$ para V. Si $\mathbf{x} = r_1\mathbf{b}_1 + \cdots + r_n\mathbf{b}_n$, entonces

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

y

$$T(\mathbf{x}) = T(r_1\mathbf{b}_1 + \cdots + r_n\mathbf{b}_n) = r_1T(\mathbf{b}_1) + \cdots + r_nT(\mathbf{b}_n) \tag{1}$$

porque T es lineal. Al usar la base $\mathcal{C}$ en W, es posible reescribir (1) en términos de vectores de $\mathcal{C}$-coordenadas:

$$[\,T(\mathbf{x})\,]_{\mathcal{C}} = r_1[\,T(\mathbf{b}_1)\,]_{\mathcal{C}} + \cdots + r_n[\,T(\mathbf{b}_n)\,]_{\mathcal{C}} \tag{2}$$

Como los vectores de $\mathcal{C}$-coordenadas están en $\mathbb{R}^m$, la ecuación vectorial (2) puede escribirse como una ecuación de matrices, a saber,

$$[T(\mathbf{x})]_{\mathcal{C}} = M[\mathbf{x}]_{\mathcal{B}} \tag{3}$$

donde

$$M = \left[\,[\,T(\mathbf{b}_1)\,]_{\mathcal{C}} \quad [\,T(\mathbf{b}_2)\,]_{\mathcal{C}} \quad \cdots \quad [\,T(\mathbf{b}_n)\,]_{\mathcal{C}}\,\right] \tag{4}$$

La matriz M es una representación matricial de T, llamada **matriz para T relativa a las bases $\mathcal{B}$ y $\mathcal{C}$**. Vea la figura 2.

La ecuación (3) postula que, en lo concerniente a los vectores de coordenadas, la acción de T sobre $\mathbf{x}$ puede verse como una multiplicación izquierda por M.

$\mathbf{x}$ —T→ $T(\mathbf{x})$
$[\mathbf{x}]_{\mathcal{B}}$ —Multiplicación por M→ $[T(\mathbf{x})]_{\mathcal{C}}$

FIGURA 2

EJEMPLO 1 Suponga que $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ es una base para V y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ es una base para W. Sea $T : V \to W$ una transformación lineal con la propiedad de que

$$T(\mathbf{b}_1) = 3\mathbf{c}_1 - 2\mathbf{c}_2 + 5\mathbf{c}_3 \quad \text{y} \quad T(\mathbf{b}_2) = 4\mathbf{c}_1 + 7\mathbf{c}_2 - \mathbf{c}_3$$

Encuentre la matriz M para T relativa a $\mathcal{B}$ y $\mathcal{C}$.

Solución Los vectores de $\mathcal{C}$-coordenadas de las *imágenes* de $\mathbf{b}_1$ y $\mathbf{b}_2$ son

$$[\,T(\mathbf{b}_1)\,]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \quad \text{y} \quad [\,T(\mathbf{b}_2)\,]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$$

Por lo tanto,

$$M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$$

Si $\mathcal{B}$ y $\mathcal{C}$ son bases para el mismo espacio V y T es la transformación identidad $T(\mathbf{x}) = \mathbf{x}$ para $\mathbf{x}$ en V, entonces la matriz M de (4) es simplemente una matriz de cambio de coordenadas (vea la sección 4.7).

Transformaciones lineales de V en V

En el caso común donde W es igual a V y la base $\mathcal{C}$ es igual a $\mathcal{B}$, la matriz M presentada en la ecuación (4) se denomina **matriz para *T* relativa a $\mathcal{B}$**, o simplemente **$\mathcal{B}$-matriz para *T***, y se denota mediante $[T]_{\mathcal{B}}$. Vea la figura 3.

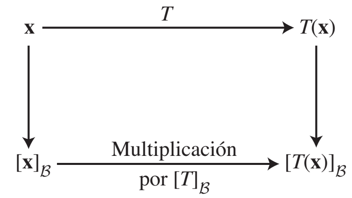


FIGURA 3

La $\mathcal{B}$-matriz de $T : V \to V$ satisface

$$[\,T(\mathbf{x})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}\,[\mathbf{x}]_{\mathcal{B}}, \quad \text{para toda } \mathbf{x} \text{ en } V \tag{5}$$

EJEMPLO 2 La función $T : \mathbb{P}_2 \to \mathbb{P}_2$ definida por

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

es una transformación lineal. (Los estudiantes de cálculo reconocerán a T como el operador de diferenciación.)

a. Encuentre la $\mathcal{B}$-matriz para T, cuando $\mathcal{B}$ es la base $\{1, t, t^2\}$.

b. Verifique si $[T(\mathbf{p})]_{\mathcal{B}} = [T]_{\mathcal{B}}\,[\mathbf{p}]_{\mathcal{B}}$ para cada $\mathbf{p}$ en $\mathbb{P}_2$.

Solución

a. Determine las imágenes de los vectores de base:

$$\begin{aligned} T(1) &= 0 && \text{El polinomio cero} \\ T(t) &= 1 && \text{El polinomio cuyo valor es siempre 1} \\ T(t^2) &= 2t \end{aligned}$$

Luego escriba los vectores de $\mathcal{B}$-coordenadas de $T(1)$, $T(t)$, y $T(t^2)$ (que en este ejemplo se encuentran mediante inspección) y colóquelos como la $\mathcal{B}$-matriz para T:

$$[T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t^2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

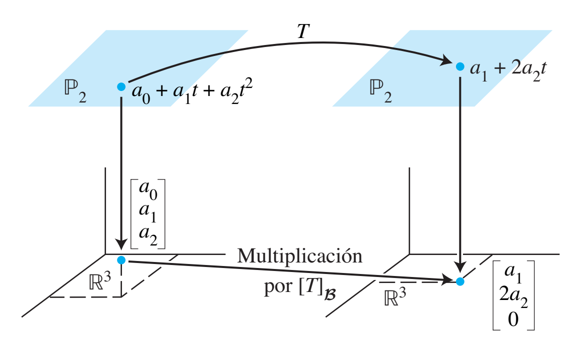
$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

b. En general, para una $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$,

$$[T(\mathbf{p})]_{\mathcal{B}} = [a_1 + 2a_2 t]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_{\mathcal{B}}[\mathbf{p}]_{\mathcal{B}}$$

Vea la figura 4.



WEB

FIGURA 4 Representación matricial de una transformación lineal.

Transformaciones lineales en $\mathbb{R}^n$

En un problema aplicado en que interviene $\mathbb{R}^n$, una transformación lineal T aparece primero, normalmente, como una transformación de matrices, $\mathbf{x} \mapsto A\mathbf{x}$. Si A es diagonalizable, entonces existe una base $\mathcal{B}$ para $\mathbb{R}^n$ que consiste en vectores propios de A. El

teorema 8 siguiente muestra que, en este caso, la $\mathcal{B}$-matriz de T es diagonal. Diagonalizar A equivale a encontrar una representación matricial diagonal de $\mathbf{x} \mapsto A\mathbf{x}$.

TEOREMA 8 **Representación de matrices diagonales**

Suponga que $A = PDP^{-1}$, donde D es una matriz diagonal de $n \times n$. Si $\mathcal{B}$ es la base para $\mathbb{R}^n$ formada a partir de las columnas de P, entonces D es la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$.

DEMOSTRACIÓN Las columnas de P se denotan mediante $\mathbf{b}_1, \dots, \mathbf{b}_n$, de manera que $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ y $P = [\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n]$. En este caso, P es la matriz $P_{\mathcal{B}}$ de cambio de coordenadas que se analizó en la sección 4.4, donde

$$P[\mathbf{x}]_{\mathcal{B}} = \mathbf{x} \quad \text{y} \quad [\mathbf{x}]_{\mathcal{B}} = P^{-1}\mathbf{x}$$

Si $T(\mathbf{x}) = A\mathbf{x}$ para $\mathbf{x}$ en $\mathbb{R}^n$, entonces

$$\begin{aligned}
[\,T\,]_{\mathcal{B}} &= \big[\,[\,T(\mathbf{b}_1)\,]_{\mathcal{B}} \quad \cdots \quad [\,T(\mathbf{b}_n)\,]_{\mathcal{B}}\,\big] && \text{Definición de } [\,T\,]_{\mathcal{B}} \\
&= \big[\,[\,A\mathbf{b}_1\,]_{\mathcal{B}} \quad \cdots \quad [\,A\mathbf{b}_n\,]_{\mathcal{B}}\,\big] && \text{Puesto que } T(\mathbf{x}) = A\mathbf{x} \\
&= [\,P^{-1}A\mathbf{b}_1 \quad \cdots \quad P^{-1}A\mathbf{b}_n\,] && \text{Cambio de coordenadas} \\
&= P^{-1}A\,[\,\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n\,] && \text{Multiplicación de matrices} \\
&= P^{-1}AP && (6)
\end{aligned}$$

Puesto que $A = PDP^{-1}$, se tiene a $[\,T\,]_{\mathcal{B}} = P^{-1}AP = D$. ■

EJEMPLO 3 Defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ como $T(\mathbf{x}) = A\mathbf{x}$, donde $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Encuentre una base $\mathcal{B}$ para $\mathbb{R}^2$ con la propiedad de que la $\mathcal{B}$-matriz de T es una matriz diagonal.

Solución A partir del ejemplo 2 de la sección 5.3, se sabe que $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

Las columnas de P, llamadas $\mathbf{b}_1$ y $\mathbf{b}_2$, son vectores propios de A. Según el teorema 8, D es la $\mathcal{B}$-matriz de T cuando $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Las funciones $\mathbf{x} \mapsto A\mathbf{x}$ y $\mathbf{u} \mapsto D\mathbf{u}$ describen la misma transformación lineal, relativa a bases diferentes. ■

Semejanza de representaciones de matrices

La demostración del teorema 8 no utilizó la información de que D era diagonal. Por lo tanto, si A es semejante a una matriz C, con $A = PCP^{-1}$, entonces C es la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ cuando la base $\mathcal{B}$ se forma a partir de las columnas de P. En la figura 5, se muestra la factorización $A = PCP^{-1}$.

De manera recíproca, si $T : \mathbb{R}^n \to \mathbb{R}^n$ está definida por medio de $T(\mathbf{x}) = A\mathbf{x}$, y si $\mathcal{B}$ es cualquier base de $\mathbb{R}^n$, entonces la $\mathcal{B}$-matriz de T es semejante a A. De hecho, los cálculos efectuados en (6) muestran que si P es la matriz cuyas columnas provienen de los vectores en $\mathcal{B}$, entonces $[T]_{\mathcal{B}} = P^{-1}AP$. Así, el conjunto de todas las matrices semejantes

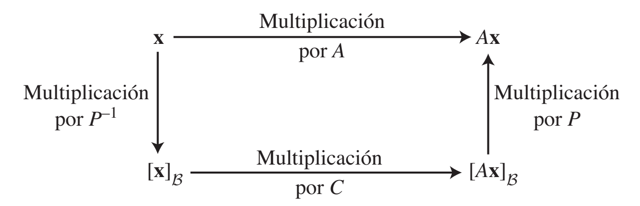


FIGURA 5 Semejanza de dos representaciones de matrices: $A = PCP^{-1}$.

a la matriz A coincide con el conjunto de todas las representaciones matriciales de la transformación $\mathbf{x} \mapsto A\mathbf{x}$.

EJEMPLO 4 Sean $A = \begin{bmatrix} 4 & -9 \\ 4 & 8 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, y $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. El polinomio característico de A es $(\lambda + 2)^2$, pero el espacio propio para el valor propio -2 es solamente unidimensional; de modo que A no es diagonalizable. Sin embargo, la base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$tiene la propiedad de que la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es una matriz triangular llamada la *forma Jordan* de A.[1] Encuentre esta $\mathcal{B}$-matriz.

Solución Si $P = [\mathbf{b}_1 \quad \mathbf{b}_2]$, entonces la $\mathcal{B}$-matriz es $P^{-1}AP$. Calcule

$$AP = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix}$$

$$P^{-1}AP = \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix}\begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}$$

Observe que el valor propio de A está sobre la diagonal.

NOTA NUMÉRICA

Una manera eficiente de calcular una $\mathcal{B}$-matriz $P^{-1}AP$ es determinar AP y después reducir por filas la matriz aumentada $[P \quad AP]$ a $[I \quad P^{-1}AP]$. Resulta innecesario calcular por separado P^{-1}. Vea el ejercicio 12 de la sección 2.2

PROBLEMAS DE PRÁCTICA

1. Encuentre $T(a_0 + a_1t + a_2t^2)$, si T es la transformación lineal de $\mathbb{P}_2$ a $\mathbb{P}_2$ cuya matriz relativa a $\mathcal{B} = \{1, t, t^2\}$ es

$$[T]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$$

[1] Cualquier matriz cuadrada A es semejante a una matriz en la forma Jordan. La base usada para producir una forma Jordan consta de vectores propios y de lo que se ha dado en llamar "vectores propios generalizados" de A. Vea el capítulo 9 de *Applied Linear Algebra*, 3a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 1988), de B. Noble y J. W. Daniel.

2. Sean A, B y C matrices de $n \times n$. El texto ha mostrado que si A es semejante a B, entonces B es semejante a A. Esta propiedad, junto con los enunciados siguientes, muestra que "semejante a" es una *relación de equivalencia.* (La equivalencia por filas es otro ejemplo de una relación de equivalencia.) Verifique los incisos (a) y (b).

a. A es semejante a A.

b. Si A es semejante a B y B es semejante a C, entonces A es semejante a C.

5.4 EJERCICIOS

1. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ bases para los espacios vectoriales V y W, respectivamente. Sea $T : V \to W$ una transformación lineal con la propiedad de que

$$T(\mathbf{b}_1) = 3\mathbf{d}_1 - 5\mathbf{d}_2, \quad T(\mathbf{b}_2) = -\mathbf{d}_1 + 6\mathbf{d}_2, \quad T(\mathbf{b}_3) = 4\mathbf{d}_2$$

Encuentre la matriz para T relativa a $\mathcal{B}$ y $\mathcal{D}$.

2. Sean $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ bases para los espacios vectoriales V y W, respectivamente. Sea $T : V \to W$ una transformación lineal con la propiedad de que

$$T(\mathbf{d}_1) = 2\mathbf{b}_1 - 3\mathbf{b}_2, \qquad T(\mathbf{d}_2) = -4\mathbf{b}_1 + 5\mathbf{b}_2$$

Encuentre la matriz para T relativa a $\mathcal{D}$ y $\mathcal{B}$.

3. Sean $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base canónica para $\mathbb{R}^3$, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V, y $T : \mathbb{R}^3 \to V$ una transformación lineal con la propiedad de que

$$T(x_1, x_2, x_3) = (x_3 - x_2)\mathbf{b}_1 - (x_1 + x_3)\mathbf{b}_2 + (x_1 - x_2)\mathbf{b}_3$$

a. Calcule $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$ y $T(\mathbf{e}_3)$.

b. Calcule $[T(\mathbf{e}_1)]_{\mathcal{B}}$, $[T(\mathbf{e}_2)]_{\mathcal{B}}$ y $[T(\mathbf{e}_3)]_{\mathcal{B}}$.

c. Encuentre la matriz para T relativa a $\mathcal{E}$ y $\mathcal{B}$.

4. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V y $T : V \to \mathbb{R}^2$ una transformación lineal con la propiedad de que

$$T(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3) = \begin{bmatrix} 2x_1 - 4x_2 + 5x_3 \\ -x_2 + 3x_3 \end{bmatrix}$$

Encuentre la matriz para T relativa a $\mathcal{B}$ y la base canónica para $\mathbb{R}^2$.

5. Sea $T : \mathbb{P}_2 \to \mathbb{P}_3$ la transformación que mapea un polinomio $\mathbf{p}(t)$ en el polinomio $(t + 5)\mathbf{p}(t)$.

a. Encuentre la imagen de $\mathbf{p}(t) = 2 - t + t^2$.

b. Muestre que T es una transformación lineal.

c. Encuentre la matriz para T relativa a las bases $\{1, t, t^2\}$ y $\{1, t, t^2, t^3\}$.

6. Sea $T : \mathbb{P}_2 \to \mathbb{P}_4$ la transformación que mapea un polinomio $\mathbf{p}(t)$ en el polinomio $\mathbf{p}(t) + t^2\mathbf{p}(t)$.

a. Encuentre la imagen de $\mathbf{p}(t) = 2 - t + t^2$.

b. Muestre que T es una transformación lineal.

c. Encuentre la matriz para T relativa a las bases $\{1, t, t^2\}$ y $\{1, t, t^2, t^3, t^4\}$.

7. Suponga que la función $T : \mathbb{P}_2 \to \mathbb{P}_2$ definida por

$$T(a_0 + a_1t + a_2t^2) = 3a_0 + (5a_0 - 2a_1)t + (4a_1 + a_2)t^2$$

es lineal. Encuentre la representación matricial de T relativa a la base $\mathcal{B} = \{1, t, t^2\}$.

8. Sea $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V. Encuentre $T(3\mathbf{b}_1 - 4\mathbf{b}_2)$ cuando T es una transformación lineal de V a V cuya matriz relativa a $\mathcal{B}$ es

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & -6 & 1 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$$

9. Defina $T : \mathbb{P}_2 \to \mathbb{R}^3$ by $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-1) \\ \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$.

a. Encuentre la imagen bajo T de $\mathbf{p}(t) = 5 + 3t$.

b. Demuestre que T es una transformación lineal.

c. Encuentre la matriz para T relativa a la base $\{1, t, t^2\}$ para $\mathbb{P}_2$ y la base estándar para $\mathbb{R}^3$.

10. Defina $T : \mathbb{P}_3 \to \mathbb{R}^4$ como $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-3) \\ \mathbf{p}(-1) \\ \mathbf{p}(1) \\ \mathbf{p}(3) \end{bmatrix}$.

a. Muestre que T es una transformación lineal.

b. Encuentre la matriz para T relativa a la base $\{1, t, t^2, t^3\}$ para $\mathbb{P}_3$ y la base canónica para $\mathbb{R}^4$.

En los ejercicios 11 y 12, encuentre la $\mathcal{B}$-matriz de la transformación $\mathbf{x} \mapsto A\mathbf{x}$, cuando $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

11. $A = \begin{bmatrix} 3 & 4 \\ -1 & -1 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

12. $A = \begin{bmatrix} -1 & 4 \\ -2 & 3 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

En los ejercicios 13 a 16, defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ como $T(\mathbf{x}) = A\mathbf{x}$. Encuentre una base $\mathcal{B}$ para $\mathbb{R}^2$ con la propiedad de que $[T]_{\mathcal{B}}$ sea diagonal.

13. $A = \begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix}$

14. $A = \begin{bmatrix} 5 & -3 \\ -7 & 1 \end{bmatrix}$

15. $A = \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}$

16. $A = \begin{bmatrix} 2 & -6 \\ -1 & 3 \end{bmatrix}$

17. Sean $A = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, para $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$. Defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ como $T(\mathbf{x}) = A\mathbf{x}$.

a. Verifique que $\mathbf{b}_1$ es un vector propio de A pero A no es diagonalizable.

b. Encuentre la $\mathcal{B}$-matriz para T.

18. Defina $T : \mathbb{R}^3 \to \mathbb{R}^3$ como $T(\mathbf{x}) = A\mathbf{x}$, donde A es una matriz de 3×3 con valores propios 5 y -2. ¿Existe una base $\mathcal{B}$ para $\mathbb{R}^3$ tal que la $\mathcal{B}$-matriz para T sea una matriz diagonal? Analice el planteamiento.

Verifique los enunciados de los ejercicios 19 a 24. Las matrices son cuadradas.

19. Si A es invertible y semejante a B, entonces B es invertible y A^{-1} es semejante a B^{-1}. [*Sugerencia:* $P^{-1}AP = B$ para alguna P invertible. Explique por qué B es invertible. Luego encuentre una Q invertible tal que $Q^{-1}A^{-1}Q = B^{-1}$.]

20. Si A es semejante a B, entonces A^2 es semejante a B^2.

21. Si B es semejante a A y C es semejante a A, entonces B es semejante a C.

22. Si A es diagonalizable y B es similar a A, entonces B también es diagonalizable.

23. Si $B = P^{-1}AP$ y $\mathbf{x}$ es un vector propio de A correspondiente a un valor propio λ, entonces $P^{-1}\mathbf{x}$ es un vector propio de B también correspondiente a λ.

24. Si A y B son semejantes, entonces tienen el mismo rango. [*Sugerencia:* Remítase a los ejercicios suplementarios 13 y 14 del capítulo 4.]

25. La *traza* de una matriz cuadrada A es la suma de las entradas diagonales en A y se denota mediante tr A. Se puede verificar que $\text{tr}(FG) = \text{tr}(GF)$ para cualesquiera dos matrices F y G de $n \times n$. Muestre que si A y B son semejantes, entonces tr A = tr B.

26. Se puede mostrar que la traza de una matriz A es igual a la suma de los valores propios de A. Verifique este enunciado para el caso en que A sea diagonalizable.

27. Sea $V = \mathbb{R}^n$ con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$; sea $W = \mathbb{R}^n$ con la base estándar, denotada aquí mediante $\mathcal{E}$; y considere la transformación identidad $I : \mathbb{R}^n \to \mathbb{R}^n$, donde $I(\mathbf{x}) = \mathbf{x}$. Encuentre la matriz para I relativa a $\mathcal{B}$ y $\mathcal{E}$. ¿Cómo se llamó a esta matriz en la sección 4.4?

28. Sean V un espacio vectorial con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, W el mismo espacio V con una base $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$, e I la transformación identidad $I : V \to W$. Encuentre la matriz para I relativa a $\mathcal{B}$ y $\mathcal{C}$. ¿Cómo se llamó a esta matriz en la sección 4.7?

29. Sea V un espacio vectorial con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Encuentre la $\mathcal{B}$-matriz de la transformación identidad $I : V \to V$.

[M] En los ejercicios 30 y 31, encuentre la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ cuando $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

30. $A = \begin{bmatrix} -14 & 4 & -14 \\ -33 & 9 & -31 \\ 11 & -4 & 11 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix},$
$\mathbf{b}_3 = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix}$

31. $A = \begin{bmatrix} -7 & -48 & -16 \\ 1 & 14 & 6 \\ -3 & -45 & -19 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix},$
$\mathbf{b}_3 = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$

32. **[M]** Sea T la transformación cuya matriz estándar se presenta enseguida. Encuentre una base para $\mathbb{R}^4$ con la propiedad de que $[T]_{\mathcal{B}}$ sea diagonal.

$$A = \begin{bmatrix} 15 & -66 & -44 & -33 \\ 0 & 13 & 21 & -15 \\ 1 & -15 & -21 & 12 \\ 2 & -18 & -22 & 8 \end{bmatrix}$$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sea $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$ y calcule

$$[\,T(\mathbf{p})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}\,[\,\mathbf{p}\,]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

Así que $T(\mathbf{p}) = (3a_0 + 4a_1) + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$.

2. a. $A = (I)^{-1}AI$, entonces A es similar a A.

b. Por hipótesis, existen matrices invertibles P y Q con la propiedad de que $B = P^{-1}AP$ y $C = Q^{-1}BQ$. Sustituya la fórmula para B en la fórmula para C, y aplique un hecho acerca del inverso de un producto:

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$$

Esta ecuación tiene la forma adecuada para demostrar que A es semejante a C.

5.5 VALORES PROPIOS COMPLEJOS

Puesto que la ecuación característica de una matriz de $n \times n$ implica un polinomio de grado n, la ecuación tiene siempre exactamente n raíces, contando multiplicidades, *a condición de que se incluyan las posibles raíces complejas.* En esta sección se muestra que si la ecuación característica de una matriz real A tiene algunas raíces complejas, entonces estas raíces proporcionan información crítica acerca de A. La clave es dejar que A actúe sobre el espacio $\mathbb{C}^n$ de n-adas de números complejos.[1]

El interés en $\mathbb{C}^n$ no surge de un deseo de "generalizar" los resultados de los capítulos anteriores, aunque eso, de hecho, abriría nuevas e importantes áreas de aplicación del álgebra lineal.[2] En vez de ello, el estudio de valores propios complejos es indispensable para descubrir información "oculta" acerca de ciertas matrices con entradas reales que surgen en diversos problemas de la vida real. Tales problemas incluyen muchos sistemas dinámicos reales en los que interviene un movimiento periódico, una vibración, o algún tipo de rotación en el espacio.

La teoría de valor propio-vector propio de matrices ya desarrollada para $\mathbb{R}^n$ se aplica igualmente bien a $\mathbb{C}^n$. De manera que un escalar complejo λ satisface $\det(A - \lambda I) = 0$ si, y sólo si, hay un vector $\mathbf{x}$ diferente de cero en $\mathbb{C}^n$ tal que $A\mathbf{x} = \lambda\mathbf{x}$. A λ se le denomina **valor propio (complejo)** y a $\mathbf{x}$ **vector propio (complejo)** correspondiente a λ.

EJEMPLO 1 Si $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ sobre $\mathbb{R}^2$ gira el plano en sentido antihorario un cuarto de vuelta. La acción de A es periódica puesto que, después de cuatro cuartos de vuelta, un vector vuelve a donde empezó. Desde luego, ningún vector diferente de cero se mapea en un múltiplo de sí mismo, así

---

[1]Para un breve análisis de los números complejos, remítase al apéndice B. El álgebra de matrices y los conceptos que se estudian en relación con los espacios vectoriales reales se extienden al caso con entradas complejas y escalares. En particular, $A(c\mathbf{x} + d\mathbf{y}) = cA\mathbf{x} + dA\mathbf{y}$, para una matriz A *de* $m \times n$ con entradas complejas $\mathbf{x}$, $\mathbf{y}$ en $\mathbb{C}^n$, y c, d en $\mathbb{C}$.

[2]Un segundo curso de álgebra lineal a menudo analiza tales temas, que son de particular importancia en ingeniería eléctrica.

A no tiene vectores propios en $\mathbb{R}^2$ y, por lo tanto, ningún valor propio real. De hecho, la ecuación característica de A es

$$\lambda^2 + 1 = 0$$

Las únicas raíces son complejas: $\lambda = i$ y $\lambda = -i$. Sin embargo, si se permite que A actúe sobre $\mathbb{C}^2$, entonces

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} 1 \\ -i \end{bmatrix}$$
$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i\begin{bmatrix} 1 \\ i \end{bmatrix}$$

Entonces i y $-i$ son valores propios, con $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ y $\begin{bmatrix} 1 \\ i \end{bmatrix}$ como vectores propios correspondientes. (En el ejemplo 2 se analiza un método para *encontrar* vectores propios complejos.) ■

Esta sección se enfocará principalmente en la matriz del siguiente ejemplo.

EJEMPLO 2 Sea $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$. Encuentre los valores propios de A y una base para cada espacio propio.

Solución La ecuación característica de A es

$$0 = \det\begin{bmatrix} .5 - \lambda & -.6 \\ .75 & 1.1 - \lambda \end{bmatrix} = (.5 - \lambda)(1.1 - \lambda) - (-.6)(.75)$$
$$= \lambda^2 - 1.6\lambda + 1$$

De la fórmula cuadrática, $\lambda = \frac{1}{2}[1.6 \pm \sqrt{(-1.6)^2 - 4}] = .8 \pm .6i$. Para el valor propio $\lambda = .8 - .6i$, construya

$$A - (.8 - .6i)I = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} - \begin{bmatrix} .8 - .6i & 0 \\ 0 & .8 - .6i \end{bmatrix}$$
$$= \begin{bmatrix} -.3 + .6i & -.6 \\ .75 & .3 + .6i \end{bmatrix} \quad (1)$$

Es bastante molesto calcular a mano la reducción por filas de la matriz aumentada usual, a causa de la aritmética de números complejos. Sin embargo, a continuación se presenta una aguda observación que realmente simplifica las cosas: Puesto que $.8 - .6i$ es un vector propio, el sistema

$$\begin{aligned} (-.3 + .6i)x_1 - \quad\quad .6x_2 &= 0 \\ .75x_1 + (.3 + .6i)x_2 &= 0 \end{aligned} \quad (2)$$

tiene una solución no trivial (con x_1 y x_2 como posibles números complejos). Por lo tanto, *ambas ecuaciones de (2) determinan la misma relación entre x_1 y x_2, y cualesquiera de las ecuaciones puede usarse para expresar una variable en términos de la otra.*[3]

[3]Otra manera de ver esto es darse cuenta de que la matriz dada en (1) no es invertible, de modo que sus filas son dependientes linealmente (como vectores en $\mathbb{C}^2$) y, por lo tanto, una fila es un múltiplo (complejo) de otra.

La segunda ecuación de (2) conduce a

$$\begin{aligned} .75x_1 &= (-.3 - .6i)x_2 \\ x_1 &= (-.4 - .8i)x_2 \end{aligned}$$

Tomando $x_2 = 5$ para eliminar los decimales, se tiene que $x_1 = -2 - 4i$. Una base para el espacio propio correspondiente a $\lambda = .8 - .6i$ es

$$\mathbf{v}_1 = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$$

Cálculos análogos para $\lambda = .8 + .6i$ producen al vector propio

$$\mathbf{v}_2 = \begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix}$$

Como una forma de comprobar el trabajo, se calcula

$$A\mathbf{v}_2 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix} = \begin{bmatrix} -4 + 2i \\ 4 + 3i \end{bmatrix} = (.8 + .6i)\mathbf{v}_2$$

De modo sorprendente, la matriz A del ejemplo 2 determina una transformación $\mathbf{x} \mapsto A\mathbf{x}$ que es esencialmente una rotación. Esto se vuelve evidente cuando se grafican los puntos apropiados.

EJEMPLO 3 Una manera de ver cómo la multiplicación por la A del ejemplo 2 afecta los puntos es graficar un punto inicial arbitrario —por ejemplo, $\mathbf{x}_0 = (2, 0)$— y luego graficar imágenes sucesivas de este punto bajo repetidas multiplicaciones por A. Es decir, graficar

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} \\ \mathbf{x}_2 &= A\mathbf{x}_1 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -.4 \\ 2.4 \end{bmatrix} \\ \mathbf{x}_3 &= A\mathbf{x}_2, \ldots \end{aligned}$$

En la figura 1 de la página 338 se muestran $\mathbf{x}_0, \ldots, \mathbf{x}_8$ como puntos gruesos. Los puntos pequeños son las posiciones de $\mathbf{x}_9, \ldots, \mathbf{x}_{100}$. La sucesión está sobre una órbita elíptica.

Por supuesto, la figura 1 no explica *por qué* ocurre la rotación. El secreto de la rotación se oculta en las partes real e imaginaria de un vector propio complejo.

Partes real e imaginaria de los vectores

El conjugado complejo de un vector complejo $\mathbf{x}$ en $\mathbb{C}^n$ es el vector $\overline{\mathbf{x}}$ en $\mathbb{C}^n$ cuyas entradas son los conjugados complejos de las entradas de $\mathbf{x}$. Las **partes real** e **imaginaria** de un vector complejo $\mathbf{x}$ son los vectores Re $\mathbf{x}$ e Im $\mathbf{x}$ formados a partir de las partes real e imaginaria de las entradas de $\mathbf{x}$.

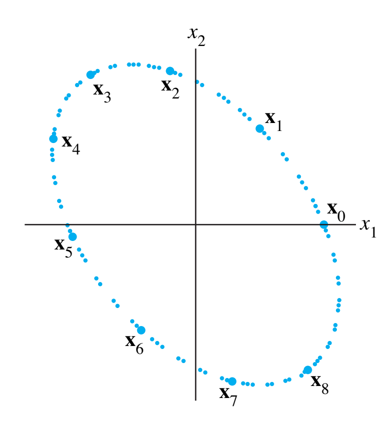


FIGURA 1 Iteraciones de un punto $\mathbf{x}_0$ bajo la acción de una matriz con un valor propio complejo.

EJEMPLO 4 Si $\mathbf{x} = \begin{bmatrix} 3-i \\ i \\ 2+5i \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}$, entonces

$$\operatorname{Re}\mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \quad \operatorname{Im}\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}, \quad \text{y} \quad \overline{\mathbf{x}} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} - i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+i \\ -i \\ 2-5i \end{bmatrix}$$

Si B es una matriz de $m \times n$ con entradas posiblemente complejas, entonces $\overline{B}$ denota la matriz cuyas entradas son los conjugados complejos de las entradas en B. Las propiedades de los conjugados para números complejos se extienden al álgebra de matrices complejas:

$$\overline{r\mathbf{x}} = \overline{r}\,\overline{\mathbf{x}}, \quad \overline{B\mathbf{x}} = \overline{B}\,\overline{\mathbf{x}}, \quad \overline{BC} = \overline{B}\,\overline{C}, \quad \text{y} \quad \overline{rB} = \overline{r}\,\overline{B}$$

Valores propios y vectores propios de una matriz real que actúa sobre $\mathbb{C}^n$

Sea A una matriz de $n \times n$ cuyas entradas son reales. Entonces $\overline{A\mathbf{x}} = \overline{A}\,\overline{\mathbf{x}} = A\overline{\mathbf{x}}$. Si λ es un valor propio de A y $\mathbf{x}$ el vector propio correspondiente en $\mathbb{C}^n$, entonces

$$A\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\,\overline{\mathbf{x}}$$

De aquí que $\overline{\lambda}$ también es un valor propio de A, con $\overline{\mathbf{x}}$ como el correspondiente vector propio. Esto muestra que *cuando A es real, sus valores propios complejos ocurren en pares conjugados.* (Aquí y en cualquier otro sitio, se utiliza el término *valor propio complejo* para hacer referencia a un valor propio $\lambda = a + bi$, con $b \neq 0$.)

EJEMPLO 5 Los valores propios de la matriz real del ejemplo 2 son conjugados complejos, a saber, $.8 - .6i$ y $.8 + .6i$. Los vectores propios correspondientes encontrados en

el ejemplo 2 también son conjugados:

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \overline{\mathbf{v}}_1$$

El siguiente ejemplo proporciona el “bloque de construcción” básico para todas las matrices reales de 2×2 con valores propios complejos.

EJEMPLO 6 Si $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, donde a y b son reales y no son ambos cero, entonces los valores propios de C son $\lambda = a \pm bi$. (Vea el problema de práctica al final de esta sección.) También, si $r = |\lambda| = \sqrt{a^2+b^2}$, entonces

$$C = r\begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}$$

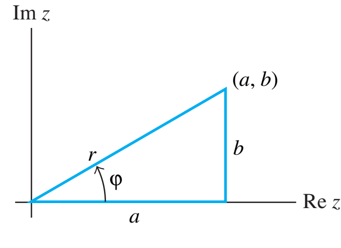


FIGURA 2

donde φ es el ángulo ubicado entre el eje x positivo y el rayo que parte de (0, 0) hacia (a, b). Vea la figura 2 y el apéndice B. El ángulo φ es el *argumento* de $\lambda = a + bi$. Entonces la transformación $\mathbf{x} \mapsto C\mathbf{x}$ puede verse como la composición de una rotación a través de un ángulo φ y la aplicación de una escala por $|\lambda|$ (vea la figura 3).

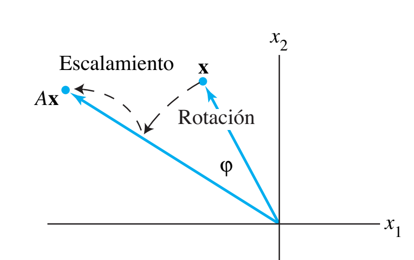


FIGURA 3 Una rotación seguida de la aplicación de una escala.

Por último, se tiene la capacidad de descubrir la rotación oculta dentro de una matriz real que tiene un valor propio complejo.

EJEMPLO 7 Sean $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$, $\lambda = .8 - .6i$, y $\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$, como en el ejemplo 2. También, sea P la matriz real de 2×2

$$P = \begin{bmatrix} \operatorname{Re}\mathbf{v}_1 & \operatorname{Im}\mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$$

y sea

$$C = P^{-1}AP = \frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}\begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix}$$

De acuerdo con el ejemplo 6, C es una rotación pura porque $|\lambda|^2 = (.8)^2 + (.6)^2 = 1$. A partir de $C = P^{-1}AP$, se obtiene

$$A = PCP^{-1} = P\begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix}P^{-1}$$

¡Ésta es la rotación que estaba "dentro" de A! La matriz P proporciona un cambio de variable, por ejemplo, $\mathbf{x} = P\mathbf{u}$. La acción de A equivale a un cambio de variable de $\mathbf{x}$ a $\mathbf{u}$, seguido de una rotación, y luego de un regreso a la variable original. Vea la figura 4. La rotación produce una elipse, como en la figura 1, en lugar de un círculo, porque el sistema de coordenadas determinado por las columnas de P no es rectangular y no tiene unidades de longitud iguales en los dos ejes.

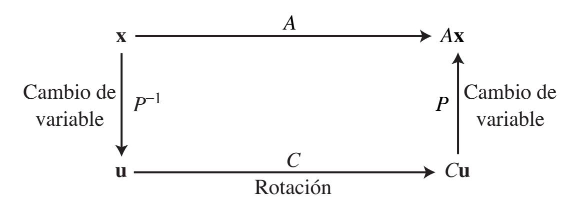


FIGURA 4 Rotación debida a un valor propio complejo.

El teorema siguiente muestra que los cálculos del ejemplo 7 pueden realizarse para cualquier matriz real A de 2×2 que tenga un valor propio complejo λ. La demostración aplica el hecho de que si las entradas de A son reales, entonces $A(\text{Re}\,\mathbf{x}) = \text{Re}\,A\mathbf{x}$ y $A(\text{Im}\,\mathbf{x}) = \text{Im}\,A\mathbf{x}$, y si $\mathbf{x}$ es un vector propio para un valor propio complejo, entonces Re $\mathbf{x}$ e Im $\mathbf{x}$ son linealmente independientes en $\mathbb{R}^2$. (Vea los ejercicios 25 y 26.) Se omiten los detalles.

TEOREMA 9

Sea A una matriz real de 2×2 con valores propios complejos $\lambda = a - bi$ $(b \neq 0)$ y un vector propio asociado $\mathbf{v}$ en $\mathbb{C}^2$. Entonces

$$A = PCP^{-1}, \quad \text{donde} \quad P = [\,\text{Re}\,\mathbf{v} \quad \text{Im}\,\mathbf{v}\,] \quad \text{y} \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

El fenómeno que se mostró en el ejemplo 7 persiste en dimensiones más altas. Por ejemplo, si A es una matriz de 3×3 con un valor propio complejo, entonces existe un plano en $\mathbb{R}^3$ sobre el cual A funciona como una rotación (posiblemente combinada con un escalamiento). Cada vector en ese plano se gira hasta otro punto ubicado sobre el mismo plano. Se afirma que el plano es **invariante** bajo A.

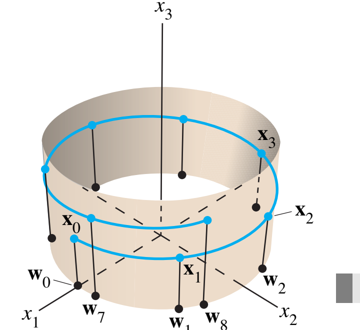


FIGURA 5 Iteraciones de dos puntos bajo la acción de una matriz de 3×3 con un valor propio complejo.

EJEMPLO 8 La matriz $A = \begin{bmatrix} .8 & -.6 & 0 \\ .6 & .8 & 0 \\ 0 & 0 & 1.07 \end{bmatrix}$ tiene valores propios $.8 \pm .6i$ y 1.07. Cualquier vector $\mathbf{w}_0$ en el plano x_1x_2 (con la tercera coordenada 0) es girado mediante A hasta otro punto localizado en el plano. Para cualquier vector $\mathbf{x}_0$ que no esté en el plano, su coordenada x_3 se multiplica por 1.07. En la figura 5, se muestran las iteraciones de los puntos $\mathbf{w}_0 = (2, 0, 0)$ y $\mathbf{x}_0 = (2, 0, 1)$ bajo la multiplicación por A.

PROBLEMA DE PRÁCTICA

Muestre que si a y b son reales, entonces los valores propios de $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ son $a \pm bi$, siendo los vectores propios correspondientes $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ y $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

5.5 EJERCICIOS

Cada matriz de los ejercicios 1 a 6 actúa sobre $\mathbb{C}^2$. Encuentre los valores propios y una base para cada espacio propio en $\mathbb{C}^2$.

1. $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$ **2.** $\begin{bmatrix} 5 & -5 \\ 1 & 1 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 5 \\ -2 & 3 \end{bmatrix}$ **4.** $\begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$

5. $\begin{bmatrix} 0 & 1 \\ -8 & 4 \end{bmatrix}$ **6.** $\begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix}$

En los ejercicios 7 a 12, utilice el ejemplo 6 para enlistar los valores de A. En cada caso, la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es la composición de una rotación seguida de un escalamiento. Proporcione el ángulo de rotación φ, donde $-\pi < \varphi \leq \pi$, y proporcione el factor de escala r.

7. $\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$ **8.** $\begin{bmatrix} \sqrt{3} & 3 \\ -3 & \sqrt{3} \end{bmatrix}$

9. $\begin{bmatrix} -\sqrt{3}/2 & 1/2 \\ -1/2 & -\sqrt{3}/2 \end{bmatrix}$ **10.** $\begin{bmatrix} -5 & -5 \\ 5 & -5 \end{bmatrix}$

11. $\begin{bmatrix} .1 & .1 \\ -.1 & .1 \end{bmatrix}$ **12.** $\begin{bmatrix} 0 & .3 \\ -.3 & 0 \end{bmatrix}$

En los ejercicios 13 a 20, encuentre una matriz P invertible y una matriz C de la forma $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ tales que la matriz dada tenga la forma $A = PCP^{-1}$. Para los ejercicios 13 a 16, utilice la información de los ejercicios 1 a 4.

13. $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 5 & -5 \\ 1 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 5 \\ -2 & 3 \end{bmatrix}$ **16.** $\begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$

17. $\begin{bmatrix} 1 & -.8 \\ 4 & -2.2 \end{bmatrix}$ **18.** $\begin{bmatrix} 1 & -1 \\ .4 & .6 \end{bmatrix}$

19. $\begin{bmatrix} 1.52 & -.7 \\ .56 & .4 \end{bmatrix}$ **20.** $\begin{bmatrix} -1.64 & -2.4 \\ 1.92 & 2.2 \end{bmatrix}$

21. En el ejemplo 2, resuelva la primera ecuación en (2) para x_2 en términos de x_1, y de ahí produzca el vector propio $\mathbf{y} = \begin{bmatrix} 2 \\ -1+2i \end{bmatrix}$ para la matriz A. Muestre que este $\mathbf{y}$ es un múltiplo (complejo) del vector $\mathbf{v}_1$ usado en el ejemplo 2.

22. Sea A una matriz compleja (o real) de $n \times n$, y sea $\mathbf{x}$ en $\mathbb{C}^n$ el vector propio correspondiente a un valor propio λ en $\mathbb{C}$. Muestre que para todo escalar complejo diferente de cero μ, el vector $\mu\mathbf{x}$ es un vector propio de A.

El capítulo 7 se concentra en matrices A con la propiedad de que $A^T = A$. Los ejercicios 23 y 24 muestran que todo valor propio de una matriz con tal propiedad es necesariamente real.

23. Sea A una matriz real de $n \times n$ con la propiedad de que $A^T = A$, sea $\mathbf{x}$ cualquier vector en $\mathbb{C}^n$, y sea $q = \overline{\mathbf{x}}^T A\mathbf{x}$. Las igualdades siguientes muestran que q es un número real al verificar que $\overline{q} = q$. Proporcione una razón para cada paso.

$$\overline{q} \underset{(a)}{=} \overline{\overline{\mathbf{x}}^T A\mathbf{x}} \underset{(b)}{=} \mathbf{x}^T \overline{A\mathbf{x}} \underset{(c)}{=} \mathbf{x}^T A\overline{\mathbf{x}} \underset{(d)}{=} (\mathbf{x}^T A\overline{\mathbf{x}})^T \underset{(e)}{=} \overline{\mathbf{x}}^T A^T \mathbf{x} = q$$

24. Sea A una matriz real de $n \times n$ con la propiedad de que $A^T = A$. Muestre que si $A\mathbf{x} = \lambda\mathbf{x}$ para algún vector $\mathbf{x}$ en $\mathbb{C}^n$ diferente de cero, entonces, de hecho, λ es real y la parte real de $\mathbf{x}$ es un vector propio de A. [*Sugerencia:* Calcule $\overline{\mathbf{x}}^T A\mathbf{x}$, y use el ejercicio 23. También, examine las partes real e imaginaria de $A\mathbf{x}$.]

25. Sean A una matriz real de $n \times n$ y $\mathbf{x}$ un vector en $\mathbb{C}^n$. Muestre que $\text{Re}(A\mathbf{x}) = A(\text{Re}\,\mathbf{x})$ e $\text{Im}(A\mathbf{x}) = A(\text{Im}\,\mathbf{x})$.

26. Sea A una matriz real de 2×2 con un valor propio complejo $\lambda = a - bi$ ($b \neq 0$) y un vector propio asociado $\mathbf{v}$ en $\mathbb{C}^2$.

a. Muestre que $A(\text{Re}\,\mathbf{v}) = a\,\text{Re}\,\mathbf{v} + b\,\text{Im}\,\mathbf{v}$ y $A(\text{Im}\,\mathbf{v}) = -b\,\text{Re}\,\mathbf{v} + a\,\text{Im}\,\mathbf{v}$. [*Sugerencia:* Escriba $\mathbf{v} = \text{Re}\,\mathbf{v} + i\,\text{Im}\,\mathbf{v}$, y calcule $A\mathbf{v}$.]

b. Pruebe que si P y C están dadas como en el teorema 9, entonces $AP = PC$.

[M] En los ejercicios 27 y 28, encuentre una factorización de la matriz dada A en la forma $A = PCP^{-1}$, donde C es una matriz diagonal en bloques con bloques de 2×2 de la forma indicada en el ejemplo 6. (Para cada par conjugado de valores propios, use las partes real e imaginaria de un vector propio en $\mathbb{C}^4$ para crear dos columnas de P.)

27. $\begin{bmatrix} .7 & 1.1 & 2.0 & 1.7 \\ -2.0 & -4.0 & -8.6 & -7.4 \\ 0 & -.5 & -1.0 & -1.0 \\ 1.0 & 2.8 & 6.0 & 5.3 \end{bmatrix}$

28. $\begin{bmatrix} -1.4 & -2.0 & -2.0 & -2.0 \\ -1.3 & -.8 & -.1 & -.6 \\ .3 & -1.9 & -1.6 & -1.4 \\ 2.0 & 3.3 & 2.3 & 2.6 \end{bmatrix}$

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Recuerde que es fácil comprobar si un vector es vector propio. No se necesita examinar la ecuación característica. Calcule

$$A\mathbf{x} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} a+bi \\ b-ai \end{bmatrix} = (a+bi)\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

Así que $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ es un vector propio correspondiente a $\lambda = a + bi$. De acuerdo con el análisis de esta sección, $\begin{bmatrix} 1 \\ i \end{bmatrix}$ debe ser un vector propio correspondiente a $\overline{\lambda} = a - bi$.

5.6 SISTEMAS DINÁMICOS DISCRETOS

Los valores propios y los vectores propios proporcionan la clave para entender el comportamiento a largo plazo, o *evolución*, de un sistema dinámico descrito mediante una *ecuación* en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Una ecuación de este tipo se usó para modelar el movimiento de la población en la sección 1.10, diversas cadenas de Markov en la sección 4.9, y la población de búhos manchados en el ejemplo introductorio de este capítulo. Los vectores $\mathbf{x}_k$ proporcionan información sobre el sistema conforme pasa el tiempo (denotado por k). En el ejemplo del búho manchado, $\mathbf{x}_k$ enlistaba la cantidad de búhos en tres categorías de edad en un tiempo k.

Las aplicaciones de esta sección se concentran en problemas ecológicos porque son más fáciles de enunciar y explicar que, por ejemplo, problemas de física o ingeniería. Sin embargo, los sistemas dinámicos surgen en muchos campos científicos. Por ejemplo, los cursos estándar de licenciatura en sistemas de control analizan diversos aspectos de los sistemas dinámicos. El moderno método de diseño del *espacio de estado* en tales cursos se apoya en gran medida en el álgebra de matrices.[1] La *respuesta de estado estable* de un sistema de control es el equivalente en ingeniería de lo que aquí se ha llamado "comportamiento a largo plazo" del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

Hasta el ejemplo 6, se supone que A es diagonalizable, con n vectores propios linealmente independientes, $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y valores propios correspondientes, $\lambda_1, \ldots, \lambda_n$.

[1]Vea G. F. Franklin, J. D. Powell, y A. Emami-Naeimi, *Feedback Control of Dynamic Systems*, 4a. ed. (Upper Saddle River, NJ: Prentice-Hall, 2001). Este texto de licenciatura tiene una buena introducción a los modelos dinámicos (capítulo 2). El diseño de espacio de estados se cubre en los capítulos 6 y 8.

Por conveniencia, suponga que los vectores propios están ordenados de manera que $|\lambda_1| \geq |\lambda_2| \cdots \geq |\lambda_n|$. Puesto que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base para $\mathbb{R}^n$, cualquier valor inicial $\mathbf{x}_0$ puede escribirse de forma única como

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \tag{1}$$

Esta *descomposición de vectores propios* de $\mathbf{x}_0$ determina lo que pasa con la sucesión $\{\mathbf{x}_k\}$. El siguiente cálculo generaliza el sencillo caso analizado en el ejemplo 5 de la sección 5.2. Puesto que los $\mathbf{v}_i$ son vectores propios,

$$\begin{aligned} \mathbf{x}_1 = A\mathbf{x}_0 &= c_1A\mathbf{v}_1 + \cdots + c_nA\mathbf{v}_n \\ &= c_1\lambda_1\mathbf{v}_1 + \cdots + c_n\lambda_n\mathbf{v}_n \end{aligned}$$

En general,

$$\mathbf{x}_k = c_1(\lambda_1)^k\mathbf{v}_1 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n \qquad (k = 0, 1, 2, \ldots) \tag{2}$$

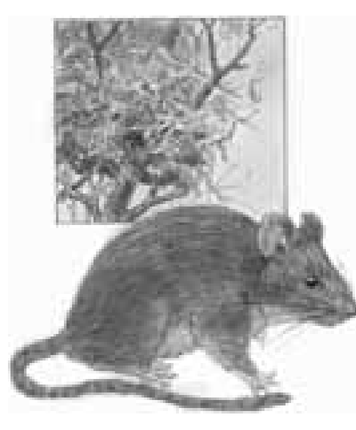
Los ejemplos siguientes ilustran lo que puede suceder en (2) como $k \to \infty$.

Un sistema depredador-presa

En California, EUA, en lo profundo de los bosques de secuoyas, las ratas pie pardo constituyen hasta el 80% de la dieta de los búhos manchados, el depredador principal de la rata de bosque. En el ejemplo 1 se utiliza un sistema lineal dinámico para modelar el sistema físico de los búhos y las ratas. (Se admite que el modelo no es realista en varios aspectos, pero puede proporcionar un punto de partida para estudiar modelos no lineales más complicados utilizados por científicos ambientalistas.)

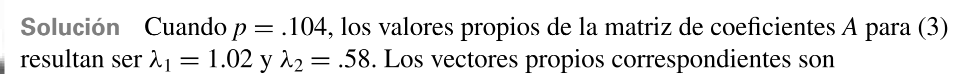
EJEMPLO 1 Denote a la población de búhos y ratas de bosque en el tiempo k mediante $\mathbf{x}_k = \begin{bmatrix} O_k \\ R_k \end{bmatrix}$, donde k es el tiempo en meses, O_k es la cantidad de búhos presentes en la región estudiada, y R_k la cantidad de ratas (medidas en miles). Suponga que

$$\begin{aligned} O_{k+1} &= (.5)O_k + (.4)R_k \\ R_{k+1} &= -p \cdot O_k + (1.1)R_k \end{aligned} \tag{3}$$



donde p es un parámetro positivo por especificar. El $(.5)O_k$ de la primera ecuación establece que sin ratas de bosque para alimentarse, sólo sobrevivirá la mitad de los búhos cada mes, mientras el $(1.1)R_k$ de la segunda ecuación señala que sin búhos como depredadores, la población de ratas aumentará en un 10% cada mes. Si hay abundancia de ratas, el $(.4)R_k$ tenderá a propiciar un aumento en la población de búhos, mientras que el término negativo $-p \cdot O_k$ mide las muertes de ratas debidas a la depredación de los búhos. (De hecho, $1000p$ es la cantidad promedio de ratas que un búho come en un mes.) Determine la evolución de este sistema cuando el parámetro de depredación p es .104.

Solución Cuando $p = .104$, los valores propios de la matriz de coeficientes A para (3) resultan ser $\lambda_1 = 1.02$ y $\lambda_2 = .58$. Los vectores propios correspondientes son



$$\mathbf{v}_1 = \begin{bmatrix} 10 \\ 13 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

Se puede escribir una $\mathbf{x}_0$ inicial como $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Entonces, para $k \geq 0$,

$$\begin{aligned}\mathbf{x}_k &= c_1(1.02)^k\mathbf{v}_1 + c_2(.58)^k\mathbf{v}_2 \\ &= c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix} + c_2(.58)^k\begin{bmatrix}5\\1\end{bmatrix}\end{aligned}$$

Cuando $k \to \infty$, $(.58)^k$ se aproxima rápidamente a cero. Suponga que $c_1 > 0$. Entonces, para toda k suficientemente grande, $\mathbf{x}_k$ es aproximadamente igual a $c_1(1.02)^k\mathbf{v}_1$, y se escribe

$$\mathbf{x}_k \approx c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix} \tag{4}$$

La aproximación en (4) mejora al aumentar k, así que para una k grande:

$$\mathbf{x}_{k+1} \approx c_1(1.02)^{k+1}\begin{bmatrix}10\\13\end{bmatrix} = (1.02)c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix} \approx 1.02\mathbf{x}_k \tag{5}$$

La aproximación en (5) establece que tarde o temprano ambas entradas de $\mathbf{x}_k$ (las cantidades de búhos y ratas) aumentarán cada mes por un factor de casi 1.02, un índice de crecimiento mensual del 2%. Según (4), $\mathbf{x}_k$ es aproximadamente un múltiplo de (10, 13), de este modo, las entradas en $\mathbf{x}_k$ tienen casi la misma proporción que 10 a 13. Esto es, por cada 10 búhos hay aproximadamente 13 mil ratas.

En el ejemplo 1 se ilustran dos afirmaciones generales acerca de un sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ en el cual A es de $n \times n$, sus valores propios satisfacen $|\lambda_1| \geq 1$ y $1 > |\lambda_j|$ para $j = 2, \ldots, n$, y $\mathbf{v}_1$ es un vector propio correspondiente a λ_1. Si $\mathbf{x}_0$ está dado por (1), con $c_1 \neq 0$, entonces para toda k suficientemente grande,

$$\mathbf{x}_{k+1} \approx \lambda_1\mathbf{x}_k \tag{6}$$

y

$$\mathbf{x}_k \approx c_1(\lambda_1)^k\mathbf{v}_1 \tag{7}$$

Las aproximaciones en (6) y (7) pueden hacerse tan cercanas como se desee al tomar una k suficientemente grande. De acuerdo con (6), $\mathbf{x}_k$ finalmente crece en un factor de casi λ_1 cada vez, así que λ_1 determina la tasa de crecimiento final del sistema. También, según (7) la proporción de cualesquiera dos entradas en $\mathbf{x}_k$ (para k grande) es casi igual a la razón de las entradas correspondientes en $\mathbf{v}_1$. El caso en que $\lambda_1 = 1$ se ilustra en el ejemplo 5 de la sección 5.2.

Descripción gráfica de soluciones

Cuando A es de 2×2, se pueden complementar los cálculos algebraicos con una descripción geométrica de la evolución del sistema. Es posible ver la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ como una descripción de lo que le sucede a un punto inicial $\mathbf{x}_0$ en $\mathbb{R}^2$ al ser transformado repetidas veces por la función $\mathbf{x} \mapsto A\mathbf{x}$. La gráfica de $\mathbf{x}_0, \mathbf{x}_1, \ldots$ es una **trayectoria** del sistema dinámico.

EJEMPLO 2 Grafique varias trayectorias del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, cuando

$$A = \begin{bmatrix}.80 & 0\\0 & .64\end{bmatrix}$$

Solución Los valores propios de A son .8 y .64, con vectores propios $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Si $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, entonces

$$\mathbf{x}_k = c_1(.8)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.64)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Por supuesto, $\mathbf{x}_k$ tiende a $\mathbf{0}$ porque tanto $(.8)^k$ como $(.64)^k$ se aproximan a 0 cuando $k \to \infty$. Pero el *camino* recorrido por $\mathbf{x}_k$ hacia $\mathbf{0}$ resulta interesante. La figura 1 muestra los primeros términos de varias trayectorias que comienzan en algunos puntos situados en los límites de la caja cuyas esquinas están en $(\pm 3, \pm 3)$. Los puntos de cada trayectoria están conectados mediante una curva delgada, para que la trayectoria sea más fácil de apreciar.

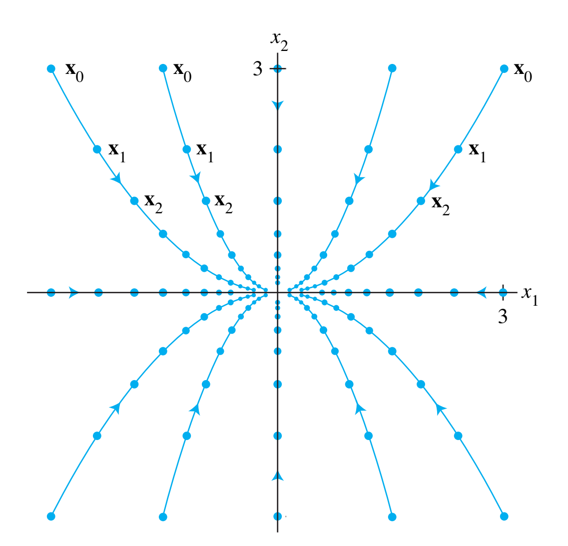


FIGURA 1 El origen como atractor.

En el ejemplo 2, al origen se le llama **atractor** del sistema dinámico porque todas las trayectorias tienden hacia $\mathbf{0}$. Esto ocurre siempre que ambos valores propios tienen magnitud menor que 1. El sentido de mayor atracción se sitúa a lo largo de la línea que pasa por $\mathbf{0}$ y el vector propio $\mathbf{v}_2$ para el valor propio de menor magnitud.

En el ejemplo siguiente, ambos valores propios de A tienen magnitud mayor que 1, y se afirma que $\mathbf{0}$ es un **repulsor** del sistema dinámico. Todas las soluciones de $\mathbf{x}_{k+1} = A\mathbf{x}_k$, excepto la solución (constante) cero, no tienen límite y tienden a alejarse del origen.[2]

[2]El origen es el único atractor o repulsor posible en un sistema dinámico *lineal*, pero puede haber múltiples atractores y repulsores en un sistema dinámico más general para el cual la función $\mathbf{x}_k \mapsto \mathbf{x}_{k+1}$ es no lineal. En un sistema de este tipo, los atractores y repulsores se definen en términos de los valores propios de una matriz especial (con entradas variables) llamada *matriz jacobiana* del sistema.

EJEMPLO 3 Grafique varias soluciones típicas de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde

$$A = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.2 \end{bmatrix}$$

Solución Los valores propios de A son 1.44 y 1.2. Si $\mathbf{x}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, entonces

$$\mathbf{x}_k = c_1(1.44)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(1.2)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Ambos términos aumentan de tamaño, pero el primero crece más rápido. Así que el sentido de mayor repulsión es la línea que pasa por **0** y por el vector propio con el valor propio de mayor magnitud. En la figura 2 se muestran varias trayectorias que empiezan en puntos muy cercanos a **0**.

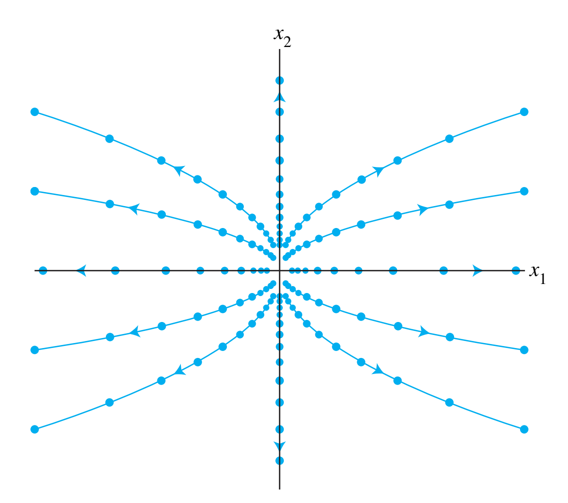


FIGURA 2 El origen como repulsor.

En el ejemplo siguiente, a **0** se le conoce como **punto silla** porque el origen atrae soluciones desde ciertos sentidos y las repele en otras direcciones. Esto ocurre siempre que un valor propio tiene magnitud mayor que 1 y otro valor tiene magnitud menor que 1. El sentido de mayor atracción está determinado por un vector propio para el valor propio de menor magnitud. El sentido de mayor repulsión está determinado por un vector propio para el valor propio de mayor magnitud.

EJEMPLO 4 Grafique varias soluciones típicas de la ecuación $\mathbf{y}_{k+1} = D\mathbf{y}_k$, donde

$$D = \begin{bmatrix} 2.0 & 0 \\ 0 & 0.5 \end{bmatrix}$$

(Aquí se escribe D e **y** en vez de A y **x** porque este ejemplo se usará posteriormente.) Muestre que una solución $\{\mathbf{y}_k\}$ es no acatada si su punto inicial no está en el eje x_2.

Solución Los valores propios de D son 2 y .5. Si $\mathbf{y}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, entonces

$$\mathbf{y}_k = c_1 2^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.5)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{8}$$

Si $\mathbf{y}_0$ está en el eje x_2, entonces $c_1 = 0$ y $\mathbf{y}_k \to \mathbf{0}$ conforme $k \to \infty$. Pero si $\mathbf{y}_0$ no está en el eje x_2, entonces el primer término de la suma para $\mathbf{y}_k$ se vuelve arbitrariamente grande, y entonces $\{\mathbf{y}_k\}$ no está acotada. En la figura 3 se muestran diez trayectorias que comienzan cerca de, o en, el eje x_2.

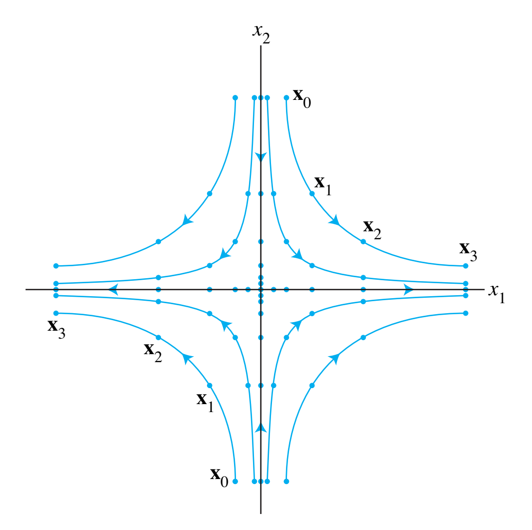


FIGURA 3 El origen como punto silla.

Cambio de variable

Los tres ejemplos precedentes se relacionaron con matrices diagonales. Para manejar el caso no diagonal, se regresará por un momento al caso $n \times n$ donde los vectores propios de A forman una base $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ para $\mathbb{R}^n$. Sea $P = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$, y sea D la matriz diagonal con los valores propios correspondientes sobre la diagonal. Dada una sucesión $\{\mathbf{x}_k\}$ que satisface $\mathbf{x}_{k+1} = A\mathbf{x}_k$, defina una nueva sucesión $\{\mathbf{y}_k\}$ mediante

$$\mathbf{y}_k = P^{-1}\mathbf{x}_k, \quad \text{o, de manera equivalente,} \quad \mathbf{x}_k = P\mathbf{y}_k$$

Al sustituir estas relaciones en la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ y usar el hecho de que PDP^{-1}, se encuentra que

$$P\mathbf{y}_{k+1} = AP\mathbf{y}_k = (PDP^{-1})P\mathbf{y}_k = PD\mathbf{y}_k$$

Si se multiplican por la izquierda ambos miembros por P^{-1}, se obtiene

$$\mathbf{y}_{k+1} = D\mathbf{y}_k$$

Al escribir $\mathbf{y}_k$ en la forma $\mathbf{y}(k)$ y denotar las entradas de $\mathbf{y}(k)$ mediante $y_1(k), \ldots, y_n(k)$, entonces

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ \vdots \\ y_n(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_n(k) \end{bmatrix}$$

El cambio de variable de $\mathbf{x}_k$ a $\mathbf{y}_k$ ha *desacoplado* el sistema de ecuaciones en diferencias. La evolución de $y_1(k)$, por ejemplo, no se ve afectada por lo que pasa con $y_2(k), \ldots, y_n(k)$, porque $y_1(k+1) = \lambda_1 \cdot y_1(k)$ para toda k.

La ecuación $\mathbf{x}_k = P\mathbf{y}_k$ establece que $\mathbf{y}_k$ es el vector de coordenadas de $\mathbf{x}_k$ con respecto a la base de vectores propios $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Es posible desacoplar el sistema $\mathbf{x}_{k+1} = A\mathbf{x}_k$ efectuando los cálculos en el nuevo sistema de coordenadas de vectores propios. Cuando $n = 2$, esto equivale a usar papel para gráficas con ejes que van en el sentido de los dos vectores propios.

EJEMPLO 5 Muestre que el origen es un punto silla para las soluciones de $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde

$$A = \begin{bmatrix} 1.25 & -.75 \\ -.75 & 1.25 \end{bmatrix}$$

Encuentre los sentidos de mayor atracción y mayor repulsión.

Solución Si se utilizan técnicas estándar, se encuentra que A tiene valores propios 2 y .5, con vectores propios correspondientes $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectivamente. Puesto que $|2| > 1$ y $|.5| < 1$, el origen es un punto silla del sistema dinámico. Si $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, entonces

$$\mathbf{x}_k = c_1 2^k \mathbf{v}_1 + c_2(.5)^k \mathbf{v}_2 \tag{9}$$

Esta ecuación parece igual a la (8) del ejemplo 4, con $\mathbf{v}_1$ y $\mathbf{v}_2$ en lugar de la base estándar.

En papel para gráficas, trace ejes que pasen por $\mathbf{0}$ y los vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$. Vea la figura 4. El movimiento a lo largo de estos ejes corresponde al movimiento trazado a lo largo de los ejes estándar en la figura 3. En la figura 4, la dirección de mayor *repulsión* es la línea que pasa por $\mathbf{0}$ y el vector propio $\mathbf{v}_1$ cuyo valor propio es mayor en magnitud que 1. Si $\mathbf{x}_0$ está sobre esta línea, el c_2 en (9) es cero y $\mathbf{x}_k$ se aleja rápidamente de $\mathbf{0}$. La dirección de mayor *atracción* está determinada por el vector propio $\mathbf{v}_2$ cuyo valor propio es menor en magnitud que 1.

En la figura 4 se muestran varias trayectorias. Cuando esta gráfica se observa en términos de los ejes de vectores propios, la imagen “parece” esencialmente la misma que la de la figura 3.

Valores propios complejos

Cuando una matriz A de 2×2 tiene valores propios complejos, A no es diagonalizable (cuando actúa sobre $\mathbb{R}^n$), pero el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ es fácil de describir. En el ejemplo 3 de la sección 5.5 se ilustra el caso en que los valores propios tienen valor absoluto de 1. Las iteraciones de un punto $\mathbf{x}_0$ están en espiral alrededor del origen a lo largo de una trayectoria elíptica.

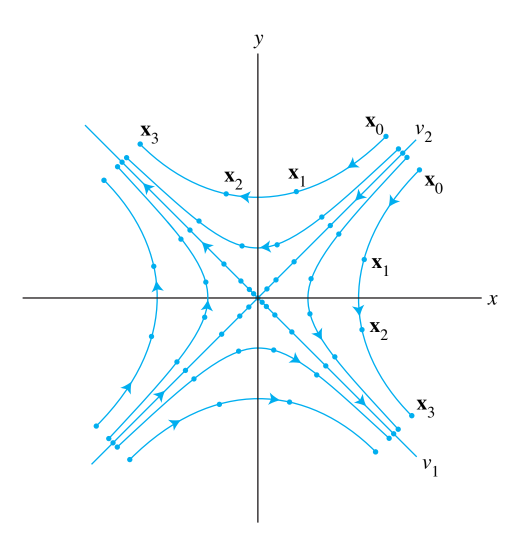


FIGURA 4 El origen como punto silla.

Si A tiene dos valores propios complejos cuyo valor absoluto es mayor que 1, entonces $\mathbf{0}$ es un repulsor y las iteraciones de $\mathbf{x}_0$ se alejarán en espiral del origen. Si los valores absolutos de los valores propios complejos son menores que 1, el origen es un atractor y las iteraciones de $\mathbf{x}_0$ se acercarán en espiral al origen, como en el ejemplo siguiente.

EJEMPLO 6 Se puede verificar que la matriz

$$A = \begin{bmatrix} .8 & .5 \\ -.1 & 1.0 \end{bmatrix}$$

tiene valores propios $.9 \pm .2i$, con vectores propios $\begin{bmatrix} 1 \mp 2i \\ 1 \end{bmatrix}$. En la figura 5 (pág. 350) se muestran tres trayectorias del sistema $\mathbf{x}_{k+1}, = A\mathbf{x}_k$ con vectores iniciales $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ -2.5 \end{bmatrix}$.

Supervivencia de los búhos manchados

Del ejemplo dado en la introducción del capítulo, recuerde que la población de búhos manchados en el área de Willow Creek en California se modeló con un sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde las entradas de $\mathbf{x}_k = (j_k, s_k, a_k)$ enlistan las cantidades de hembras (en el tiempo k) que están en las etapas de vida juvenil, subadulta y adulta, respectivamente, y A es la matriz de etapas

$$A = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \tag{10}$$

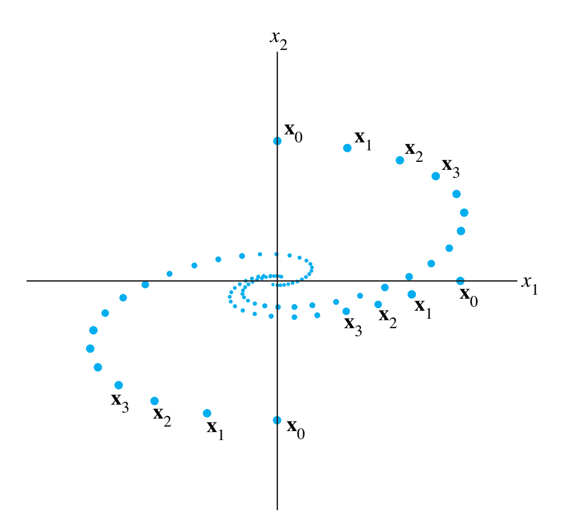


FIGURA 5 Rotación asociada con valores propios complejos.

MATLAB muestra que los valores propios de A son aproximadamente $\lambda_1 = .98$, $\lambda_2 = -.02 + .21i$, y $\lambda_3 = -.02 - .21i$. Observe que los tres valores propios tienen magnitud menor que 1, porque $|\lambda_2|^2 = |\lambda_3|^2 = (-.02)^2 + (.21)^2 = .0445$.

Por el momento, sea A tal que actúe sobre el espacio vectorial complejo $\mathbb{C}^3$. Entonces, como A tiene tres valores propios distintos, los tres vectores propios correspondientes son linealmente independientes y forman una base para $\mathbb{C}^3$. Denote los vectores propios mediante $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Por lo tanto, la solución general de $\mathbf{x}_{k+1} = A\mathbf{x}_k$ (usando vectores en $\mathbb{C}^3$) tiene la forma

$$\mathbf{x}_k = c_1(\lambda_1)^k\mathbf{v}_1 + c_2(\lambda_2)^k\mathbf{v}_2 + c_3(\lambda_3)^k\mathbf{v}_3 \tag{11}$$

Si $\mathbf{x}_0$ es un vector inicial real, entonces $\mathbf{x}_1 = A\mathbf{x}_0$ es real porque A es real. De manera similar, la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ muestra que cada $\mathbf{x}_k$ situado en la izquierda de (11) es real, aunque esté expresado como suma de vectores complejos. Sin embargo, cada término ubicado en el lado derecho de (11) se aproxima al vector cero, porque todos los valores propios tienen magnitud menor que 1. Por lo tanto, la sucesión real $\mathbf{x}_k$ también se aproxima al vector cero. Desafortunadamente, este modelo predice que todos los búhos manchados morirán tarde o temprano.

¿Hay esperanzas para el búho manchado? Recuerde, por el ejemplo introductorio, que la entrada de 18% en la matriz A en (10) proviene de que, aunque el 60% de los búhos jóvenes vive lo suficiente como para dejar el nido y buscar un nuevo territorio, sólo el 30% de ese grupo sobrevive a la búsqueda y encuentra una nueva área habitable. La cantidad de áreas boscosas taladas por completo influye de manera decisiva en la supervivencia a esta búsqueda, al volverla más difícil y peligrosa.

Parte de la población de búhos vive en áreas donde hay pocas o ningún área taladas por completo. Podría ser que allí sobreviva un porcentaje mayor de búhos juveniles y encuentre un nuevo territorio. Por supuesto, el problema de los búhos manchados es más

complejo de lo que se ha descrito aquí, pero el último ejemplo ofrece a la historia un final feliz.

EJEMPLO 7 Suponga que la tasa de supervivencia de los búhos juveniles después de la búsqueda de nuevos territorios es del 50%, de modo que la entrada (2, 1) de la matriz de etapas A en (10) es de .3 en lugar de .18. ¿Qué predice el modelo de matriz por etapas para esta población de búhos manchados?

Solución Ahora los valores propios de A son, aproximadamente, $\lambda_1 = 1.01$, $\lambda_2 = -.03 + .26i$, y $\lambda_3 = -.03 - .26i$. Un vector propio para λ_1 es, de modo aproximado, $\mathbf{v}_1 = (10, 3, 31)$. Sean $\mathbf{v}_2$ y $\mathbf{v}_3$ vectores propios (complejos) para λ_2 y λ_3. En este caso, la ecuación (11) se convierte en

$$\mathbf{x}_k = c_1(1.01)^k\mathbf{v}_1 + c_2(-.03 + .26i)^k\mathbf{v}_2 + c_3(-.03 - .26i)^k\mathbf{v}_3$$

Cuando $k \to \infty$, los segundos dos vectores tienden a cero. De manera que $\mathbf{x}_k$ se parece cada vez más al vector (real) $c_1(1.01)^k\mathbf{v}_1$. Las aproximaciones en (6) y (7), que siguen al ejemplo 1, tienen aplicación aquí. También, puede mostrarse que la constante c_1 en la descomposición inicial de $\mathbf{x}_0$ es positiva cuando las entradas en $\mathbf{x}_0$ son no negativas. Entonces la población de búhos aumentará lentamente, con una tasa de crecimiento a largo plazo de 1.01. El vector propio $\mathbf{v}_1$ describe la distribución final de los búhos por etapas vitales: Por cada 31 adultos, habrá 10 juveniles y 3 subadultos. ■

Lecturas adicionales

Franklin, G. F., J. D. Powell, y M. L. Workman, *Digital Control of Dynamic Systems*, 3a. ed. Reading, MA: Addison-Wesley, 1998.

Sandefur, James T, *Discrete Dynamical Systems—Theory and Applications*. Oxford: Oxford University Press, 1990.

Tuchinsky, Philip, *Management of a Buffalo Herd*, UMAP módulo 207. Lexington MA: COMAP, 1980.

PROBLEMAS DE PRÁCTICA

1. La siguiente matriz A tiene valores propios 1, $\frac{2}{3}$, y $\frac{1}{3}$, con vectores propios correspondientes $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$:

$$A = \frac{1}{9}\begin{bmatrix} 7 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

Encuentre la solución general de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ si $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 11 \\ -2 \end{bmatrix}$.

2. ¿Qué ocurre con la sucesión $\{\mathbf{x}_k\}$ del problema de práctica 1 conforme $k \to \infty$?

5.6 EJERCICIOS

1. Sea A una matriz de 2×2 con valores propios 3 y 1/3, y vectores propios correspondientes $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Sea $\{\mathbf{x}_k\}$ una solución de la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$, $\mathbf{x}_0 = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$.
   a. Calcule $\mathbf{x}_1 = A\mathbf{x}_0$. [*Sugerencia:* No es necesario conocer A.]
   b. Encuentre una fórmula para $\mathbf{x}_k$ que contenga a k y a los vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$.

2. Suponga que los valores propios de una matriz A de 3×3 son 3, 4/5, 3/5, con los correspondientes vectores propios $\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -5 \end{bmatrix}, \begin{bmatrix} -3 \\ -3 \\ 7 \end{bmatrix}$. Sea $\mathbf{x}_0 = \begin{bmatrix} -2 \\ -5 \\ 3 \end{bmatrix}$. Encuentre la solución de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para la $\mathbf{x}_0$ especificada, y describa qué sucede cuando $k \to \infty$.

En los ejercicios 3 a 6, suponga que cualquier vector inicial $\mathbf{x}_0$ tiene una descomposición de vectores propios tal que el coeficiente c_1 de la ecuación (1) de esta sección es positivo.[3]

3. Determine la evolución del sistema dinámico del ejemplo 1 cuando el parámetro de depredación p es de .2 en (3). (Proporcione una fórmula para $\mathbf{x}_k$.) La población de búhos aumenta o disminuye? ¿Qué sucede con la población de ratas de bosque?

4. Determine la evolución del sistema dinámico dado en el ejemplo 1 cuando el parámetro de depredación p es de .125. (Proporcione una fórmula para $\mathbf{x}_k$.) Conforme transcurre el tiempo, ¿qué sucede con el tamaño de las poblaciones de búhos y ratas? El sistema tiende hacia lo que en ocasiones se llama un equilibrio inestable. ¿Qué podría pasarle al sistema si algún aspecto del modelo (como la tasa de nacimientos o la de depredación) cambia ligeramente?

5. En los bosques de crecimiento antiguo de pinos Douglas, los búhos manchados se alimentan principalmente de ardillas voladoras. Suponga que la matriz depredador-presa para estas dos poblaciones es $A = \begin{bmatrix} .4 & .3 \\ -p & 1.2 \end{bmatrix}$. Muestre que si el parámetro de depredación p es de .325, ambas poblaciones aumentan. Estime la tasa de crecimiento a largo plazo y la proporción final de búhos a ardillas voladoras.

6. Muestre que si el parámetro de depredación dado en el ejercicio 5 es de .5, tarde o temprano morirán tanto los búhos como las ardillas. Encuentre un valor de p según el cual ambas poblaciones tiendan a mantener niveles constantes.¿Cuáles serían los tamaños de población relativa en este caso?

7. Sea A una matriz tal que tenga las propiedades descritas en el ejercicio 1.
   a. ¿El origen es un atractor, un repulsor o un punto silla del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$?
   b. Encuentre los sentidos de mayor atracción y/o repulsión para este sistema dinámico.
   c. Elabore una descripción gráfica del sistema, mostrando los sentidos de mayor atracción o repulsión. Incluya un bosquejo de varias trayectorias típicas (sin calcular puntos específicos).

8. Determine la naturaleza del origen (atractor, repulsor, punto silla) del sistema $\mathbf{x}_{k+1} = A\mathbf{x}_k$ si A tiene las propiedades descritas en el ejercicio 2. Encuentre los sentidos de mayor atracción o repulsión.

En los ejercicios 9 a 14, clasifique el origen del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ como atractor, repulsor o punto silla. Encuentre los sentidos de mayor atracción y/o repulsión.

9. $A = \begin{bmatrix} 1.7 & -.3 \\ -1.2 & .8 \end{bmatrix}$

10. $A = \begin{bmatrix} .3 & .4 \\ -.3 & 1.1 \end{bmatrix}$

11. $A = \begin{bmatrix} .4 & .5 \\ -.4 & 1.3 \end{bmatrix}$

12. $A = \begin{bmatrix} .5 & .6 \\ -.3 & 1.4 \end{bmatrix}$

13. $A = \begin{bmatrix} .8 & .3 \\ -.4 & 1.5 \end{bmatrix}$

14. $A = \begin{bmatrix} 1.7 & .6 \\ -.4 & .7 \end{bmatrix}$

15. Sea $A = \begin{bmatrix} .4 & 0 & .2 \\ .3 & .8 & .3 \\ .3 & .2 & .5 \end{bmatrix}$. El vector $\mathbf{v}_1 = \begin{bmatrix} .1 \\ .6 \\ .3 \end{bmatrix}$ es un vector propio de A, y dos valores propios son .5 y .2. Estructure la solución del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ que satisface $\mathbf{x}_0 = (0, .3, .7)$. ¿Qué le sucede a $\mathbf{x}_k$ conforme $k \to \infty$?

16. **[M]** Produzca la solución general del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ cuando A es la matriz estocástica para el modelo de Hertz Rent A Car del ejercicio 16 presentado en la sección 4.9.

17. Estructure un modelo de matriz por etapas para una especie animal que tenga dos etapas de vida: juvenil (hasta 1 año de edad) y adulta. Suponga que las hembras adultas paren una vez al año un promedio de 1.6 hembras juveniles. Cada año, sobrevive un 30% de los especímenes juveniles para convertirse en adultos, y sobrevive el 80% de los adultos. Para $k \geq 0$,

[3]Una de las limitaciones del modelo dado en el ejemplo 1 es que siempre existen vectores de población inicial $\mathbf{x}_0$ con entradas positivas, de tal forma que el coeficiente c_1 resulta negativo. La aproximación (7) aún es válida, pero las entradas en $\mathbf{x}_k$ se vuelven negativas tarde o temprano.

sea $\mathbf{x}_k = (j_k, a_k)$, donde las entradas de $\mathbf{x}_k$ son las cantidades de hembras juveniles y adultas contadas en un año k.

a. Estructure la matriz por etapas A tal que $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para $k \geq 0$.

b. Muestre que la población está aumentando, determine la tasa eventual de crecimiento de la población, y encuentre la proporción de juveniles a adultos.

c. **[M]** Suponga que inicialmente hay 15 juveniles y 10 adultos en la población. Elabore cuatro gráficas que muestren cómo cambia la población a lo largo de ocho años: (a) la cantidad de juveniles, (b) la cantidad de adultos, (c) la población total, y (d) la proporción de juveniles a adultos (cada año). ¿Cuándo parece estabilizarse la proporción en (d)? Incluya un listado del programa o las pulsaciones de tecla utilizadas para producir las gráficas en (c) y (d).

18. Se puede modelar una manada de bisontes empleando una matriz por etapas semejante a la de los búhos manchados. Las hembras pueden dividirse en becerras (hasta 1 año de edad), terneras (de 1 a 2 años), y adultas. Suponga que cada año nace un promedio de 42 becerras por cada 100 adultas. (Solamente los adultos producen crías.) Cada año sobrevive, aproximadamente, el 60% de becerras, un 75% de terneras, y el 95% de adultas. Para $k \geq 0$, sea $\mathbf{x}_k = (c_k, y_k, a_k)$, donde las entradas de x_k son la cantidad de hembras contadas en cada etapa de vida en el año k.

a. Estructure la matriz por etapas A para la manada de bisontes, tal que $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para $k \geq 0$.

b. **[M]** Muestre que la manada de bisontes está aumentando, determine la tasa de crecimiento esperada luego de muchos años, y proporcione la cantidad esperada de terneras y becerras presentes por cada 100 adultas.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. El primer paso es escribir $\mathbf{x}_0$ como una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. La reducción por filas de $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \quad \mathbf{x}_0]$ produce los pesos $c_1 = 2$, $c_2 = 1$, $c_3 = 3$, de modo que

$$\mathbf{x}_0 = 2\mathbf{v}_1 + 1\mathbf{v}_2 + 3\mathbf{v}_3$$

Puesto que los valores propios son 1, $\frac{2}{3}$, y $\frac{1}{3}$, la solución general es

$$\begin{aligned}\mathbf{x}_k &= 2 \cdot 1^k \mathbf{v}_1 + 1 \cdot \left(\frac{2}{3}\right)^k \mathbf{v}_2 + 3 \cdot \left(\frac{1}{3}\right)^k \mathbf{v}_3 \\ &= 2\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} + \left(\frac{2}{3}\right)^k \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + 3 \cdot \left(\frac{1}{3}\right)^k \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \qquad (12)\end{aligned}$$

2. Conforme $k \to \infty$, los términos segundo y tercero de (12) tienden hacia el vector cero, y

$$\mathbf{x}_k = 2\mathbf{v}_1 + \left(\frac{2}{3}\right)^k \mathbf{v}_2 + 3\left(\frac{1}{3}\right)^k \mathbf{v}_3 \to 2\mathbf{v}_1 = \begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix}$$

5.7 APLICACIONES A LAS ECUACIONES DIFERENCIALES

En esta sección se describirán análogos continuos de las ecuaciones en diferencias estudiadas en la sección 5.6. En muchos problemas aplicados, diversas cantidades cambian continuamente con el tiempo y están relacionadas mediante un sistema de ecuaciones diferenciales:

$$\begin{aligned} x_1' &= a_{11}x_1 + \cdots + a_{1n}x_n \\ x_2' &= a_{21}x_1 + \cdots + a_{2n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + \cdots + a_{nn}x_n \end{aligned}$$

Aquí $x_1, \ldots, x_n$ son funciones diferenciables de t, con derivadas $x_1', \ldots, x_n'$, y las a_{ij} son constantes. La característica crucial de este sistema es que es *lineal*. Para ver esto, escriba el sistema como una ecuación diferencial de matrices

$$\mathbf{x}' = A\mathbf{x} \tag{1}$$

donde

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \qquad \mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \qquad \text{y} \qquad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

Una **solución** de (1) es una función con valores vectoriales que satisface (1) para toda t presente en algún intervalo de números reales, como $t \geq 0$.

La ecuación (1) es *lineal* porque tanto la diferenciación de funciones como la multiplicación de vectores por una matriz son transformaciones lineales. Entonces, si $\mathbf{u}$ y $\mathbf{v}$ son soluciones de $\mathbf{x}' = A\mathbf{x}$, entonces $c\mathbf{u} + d\mathbf{v}$ también es una solución, porque

$$\begin{aligned} (c\mathbf{u} + d\mathbf{v})' &= c\mathbf{u}' + d\mathbf{v}' \\ &= cA\mathbf{u} + dA\mathbf{v} = A(c\mathbf{u} + d\mathbf{v}) \end{aligned}$$

(Los ingenieros llaman a esta propiedad *superposición* de soluciones.) De la misma forma, la función idénticamente cero es una solución (trivial) de (1). En la terminología del capítulo 4, el conjunto de todas las soluciones de (1) es un *subespacio* del conjunto de todas las funciones continuas con valores en $\mathbb{R}^n$.

Los textos estándar sobre ecuaciones diferenciales muestran que siempre existe lo que se llama un **conjunto fundamental de soluciones** de (1). Si A es de $n \times n$, entonces existen n funciones linealmente independientes en un conjunto fundamental, y cada solución de (1) es una combinación lineal única de estas n funciones. Esto es, un conjunto fundamental de soluciones es una *base* del conjunto de todas las soluciones de (1), y el conjunto solución es un espacio vectorial n-dimensional de funciones. Si un vector $\mathbf{x}_0$ se especifica, entonces el **problema con valor inicial** es estructurar la función (única) $\mathbf{x}$ tal que $\mathbf{x}' = A\mathbf{x}$ y $\mathbf{x}(0) = \mathbf{x}_0$.

Cuando A es una matriz diagonal, las soluciones de (1) pueden producirse empleando cálculo elemental. Por ejemplo, considere

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{2}$$

esto es,

$$\begin{aligned} x_1'(t) &= 3x_1(t) \\ x_2'(t) &= -5x_2(t) \end{aligned} \tag{3}$$

Se dice que el sistema (2) está *desacoplado* porque cada derivada de una función depende únicamente de la propia función, no de alguna combinación o "acoplamiento" de

$x_1(t)$ y $x_2(t)$. A partir del cálculo, se sabe que las soluciones de (3) son $x_1(t) = c_1e^{3t}$ y $x_2(t) = c_2e^{-5t}$, para cualesquiera constantes c_1 y c_2. Cada solución de (2) puede escribirse en la forma

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c_1e^{3t} \\ c_2e^{-5t} \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix}e^{3t} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}e^{-5t}$$

Este ejemplo sugiere que para la ecuación general $\mathbf{x}' = A\mathbf{x}$, una solución podría ser una combinación lineal de funciones de la forma

$$\mathbf{x}(t) = \mathbf{v}e^{\lambda t} \tag{4}$$

para algún escalar λ y algún vector fijo $\mathbf{v}$ diferente de cero. [Si $\mathbf{v} = \mathbf{0}$, la función $\mathbf{x}(t)$ es idénticamente cero y, por lo tanto, satisface $\mathbf{x}' = A\mathbf{x}$.] Observe que

$$\begin{aligned} \mathbf{x}'(t) &= \lambda\mathbf{v}e^{\lambda t} && \text{Por cálculo, puesto que } \mathbf{v} \text{ es un vector constante} \\ A\mathbf{x}(t) &= A\mathbf{v}e^{\lambda t} && \text{Multiplicando ambos lados de (4) por } A \end{aligned}$$

Puesto que $e^{\lambda t}$ nunca es cero, $\mathbf{x}'(t)$ será igual a $A\mathbf{x}(t)$ si, y sólo si, $\lambda\mathbf{v} = A\mathbf{v}$; esto es, si, y sólo si, A es un valor propio de A y $\mathbf{v}$ es el vector propio correspondiente. Entonces cada par de valor propio-vector propio proporciona una solución (4) de $\mathbf{x}' = A\mathbf{x}$. Tales soluciones algunas veces son llamadas *funciones propias* de la ecuación diferencial. Las funciones propias son clave para resolver sistemas de ecuaciones diferenciales.

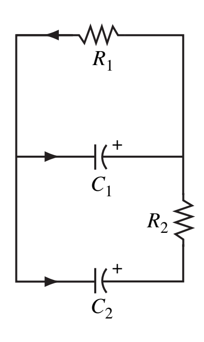


FIGURA 1

EJEMPLO 1 El circuito de la figura 1 se puede describir por medio de la ecuación diferencial

$$\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \begin{bmatrix} -(1/R_1 + 1/R_2)/C_1 & 1/(R_2C_1) \\ 1/(R_2C_2) & -1/(R_2C_2) \end{bmatrix}\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

donde $v_1(t)$ y $v_2(t)$ son los voltajes que pasan por dos condensadores en el tiempo t. Suponga que el resistor R_1 es de 1 ohm y el R_2 de 2 ohms, que el capacitor C_1 es de 1 farad y C_2 de .5 farads, y asuma una carga inicial de 5 volts en el capacitor C_1 y de 4 volts en el capacitor C_2. Encuentre las fórmulas para $v_1(t)$ y $v_2(t)$ que describan cómo cambian los voltajes con el tiempo.

Solución Para los datos dados, establezca $A = \begin{bmatrix} -1.5 & .5 \\ 1 & -1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, y $\mathbf{x}_0 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$. El vector $\mathbf{x}_0$ enlista los valores iniciales de $\mathbf{x}$. A partir de A, se obtienen los valores propios $\lambda_1 = -.5$ y $\lambda_2 = -2$, con los correspondientes vectores propios

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Las funciones propias $\mathbf{x}_1(t) = \mathbf{v}_1e^{\lambda_1 t}$ y $\mathbf{x}_2(t) = \mathbf{v}_2e^{\lambda_2 t}$ satisfacen ambas $\mathbf{x}' = A\mathbf{x}$, y también lo hace cualquier combinación lineal de $\mathbf{x}_1$ y $\mathbf{x}_2$. Sea

$$\mathbf{x}(t) = c_1\mathbf{v}_1e^{\lambda_1 t} + c_2\mathbf{v}_2e^{\lambda_2 t} = c_1\begin{bmatrix} 1 \\ 2 \end{bmatrix}e^{-.5t} + c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}e^{-2t}$$

y observe que $\mathbf{x}(0) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Puesto que resulta obvio que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes y, por lo tanto, generan $\mathbb{R}^2$, se pueden encontrar c_1 y c_2 tales que vuelvan

$$c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$
$$\uparrow \qquad\qquad \uparrow \qquad\qquad \uparrow$$
$$\mathbf{v}_1 \qquad\quad \mathbf{v}_2 \qquad\quad \mathbf{x}_0$$

$\mathbf{x}(0)$ igual a $\mathbf{x}_0$. De hecho, la ecuación conduce fácilmente a $c_1 = 3$ y $c_2 = -2$. Entonces la solución deseada de la ecuación diferencial $\mathbf{x}' = A\mathbf{x}$ es

$$\mathbf{x}(t) = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} - 2\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

o bien

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \begin{bmatrix} 3e^{-.5t} + 2e^{-2t} \\ 6e^{-.5t} - 2e^{-2t} \end{bmatrix}$$

En la figura 2 se muestra la gráfica, o *trayectoria*, de $\mathbf{x}(t)$, para $t \geq 0$, junto con las trayectorias para otros puntos iniciales. Las trayectorias de las dos funciones propias $\mathbf{x}_1$ y $\mathbf{x}_2$ están en los espacios propios de A.

Ambas funciones $\mathbf{x}_1$ y $\mathbf{x}_2$ tienden a cero conforme $t \to \infty$, pero los valores de $\mathbf{x}_2$ decaen más rápidamente porque su exponente es más negativo. Las entradas en el vector propio correspondiente $\mathbf{v}_2$ muestran que los voltajes a través de los capacitores caerán a cero tan rápidamente como sea posible si los voltajes iniciales son iguales en magnitud pero de signo opuesto.

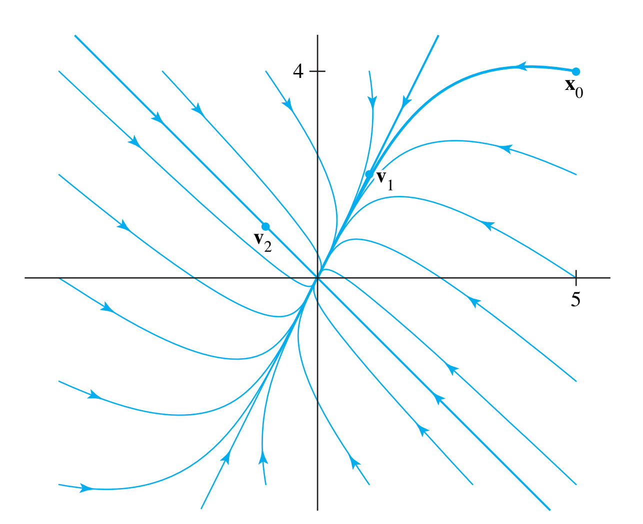


FIGURA 2 El origen como atractor.

En la figura 2, al origen se le llama **atractor**, o **sumidero**, del sistema dinámico porque todas las trayectorias son atraídas hacia el origen. La dirección de mayor atracción está a lo largo de la trayectoria de la función propia $\mathbf{x}_2$ (sobre la línea que pasa por $\mathbf{0}$ y $\mathbf{v}_2$) correspondiente al valor propio más negativo, $\lambda = -2$. Las trayectorias que comienzan en puntos que no están sobre esta línea se vuelven asintóticas a la línea que pasa por $\mathbf{0}$ y $\mathbf{v}_1$ porque sus componentes en la dirección $\mathbf{v}_2$ decaen rápidamente.

Si los valores propios del ejemplo 1 fueran positivos en vez de negativos, las trayectorias correspondientes tendrían una forma semejante, pero se recorrerían *alejándose* del origen. En un caso así, el origen recibe el nombre de **repulsor**, o **fuente**, del sistema dinámico, y la dirección de mayor repulsión es la línea que contiene la trayectoria de la función propia correspondiente al valor propio más positivo.

EJEMPLO 2 Suponga que una partícula se mueve en un campo de fuerzas plano y que su vector de posición $\mathbf{x}$ satisface $\mathbf{x}' = A\mathbf{x}$ y $\mathbf{x}(0) = \mathbf{x}_0$, donde

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 1 \end{bmatrix}, \qquad \mathbf{x}_0 = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Resuelva este problema de valor inicial, y trace la trayectoria de la partícula para $t \geq 0$.

Solución Los valores propios de A resultan ser $\lambda_1 = 6$ y $\lambda_2 = -1$, con los correspondientes vectores propios $\mathbf{v}_1 = (-5, 2)$ y $\mathbf{v}_2 = (1, 1)$. Para cualesquiera constantes c_1 y c_2, la función

$$\mathbf{x}(t) = c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t} = c_1\begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + c_2\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

es una solución de $A\mathbf{x}' = \mathbf{x}$. Se desea que c_1 y c_2 satisfagan $\mathbf{x}(0) = \mathbf{x}_0$, esto es,

$$c_1\begin{bmatrix} -5 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix} \quad \text{o bien} \quad \begin{bmatrix} -5 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Los cálculos muestran que $c_1 = -3/70$ y $c_2 = 188/70$, y entonces la función deseada es

$$\mathbf{x}(t) = \frac{-3}{70}\begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + \frac{188}{70}\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

Las trayectorias de $\mathbf{x}$ y otras soluciones se muestran en la figura 3.

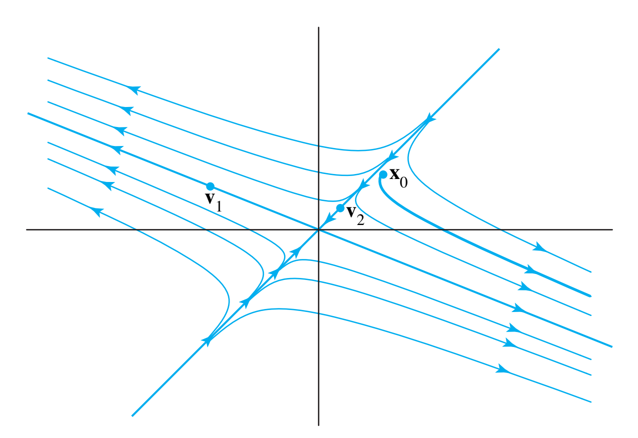


FIGURA 3 El origen como punto silla.

En la figura 3, al origen se le llama **punto silla** del sistema dinámico porque algunas trayectorias se aproximan primero al origen y luego cambian de dirección y se alejan de

él. Se presenta un punto silla siempre que la matriz A tiene valores propios tanto positivos como negativos. La dirección de mayor repulsión es la línea que pasa por $\mathbf{v}_1$ y $\mathbf{0}$, correspondiente al valor propio positivo. La dirección de mayor atracción es la línea que pasa por $\mathbf{v}_2$ y $\mathbf{0}$, correspondiente al valor propio negativo.

Desacoplamiento de un sistema dinámico

El siguiente análisis muestra que el método de los ejemplos 1 y 2 produce un conjunto fundamental de soluciones para cualquier sistema dinámico descrito por $\mathbf{x}' = A\mathbf{x}$ cuando A es de $n \times n$ y tiene n vectores propios linealmente independientes, esto es, cuando A es diagonalizable. Suponga que las funciones propias de A son

$$\mathbf{v}_1 e^{\lambda_1 t}, \quad \ldots, \quad \mathbf{v}_n e^{\lambda_n t}$$

con $\mathbf{v}_1, \ldots, \mathbf{v}_n$ como vectores propios linealmente independientes. Sea $P = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$, y sea D la matriz diagonal con entradas $\lambda_1, \ldots, \lambda_n$, de manera que $A = PDP^{-1}$. Ahora se realiza un *cambio de variable*, definiendo una nueva función $\mathbf{y}$ como

$$\mathbf{y}(t) = P^{-1}\mathbf{x}(t), \quad \text{o, de manera equivalente,} \quad \mathbf{x}(t) = P\mathbf{y}(t)$$

La ecuación $\mathbf{x}(t) = P\mathbf{y}(t)$ establece que $\mathbf{y}(t)$ es el vector de coordenadas de $\mathbf{x}(t)$ relativo a la base de vectores propios. Con la sustitución de $P\mathbf{y}$ por $\mathbf{x}$ en la ecuación $\mathbf{x}' = A\mathbf{x}$ se obtiene

$$\frac{d}{dt}(P\mathbf{y}) = A(P\mathbf{y}) = (PDP^{-1})P\mathbf{y} = PD\mathbf{y} \tag{5}$$

Puesto que P es una matriz constante, el lado izquierdo de (5) es $P\mathbf{y}'$. Multiplique por la izquierda ambos lados de (5) por P^{-1} y obtenga $\mathbf{y}' = D\mathbf{y}$, o bien

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

El cambio de variable de $\mathbf{x}$ a $\mathbf{y}$ ha *desacoplado* el sistema de ecuaciones diferenciales, porque la derivada de cada función escalar y_k depende solamente de y_k. (Revise el cambio análogo de variables de la sección 5.6.) Puesto que $y_1' = \lambda_1 y_1$, se tiene que $y_1(t) = c_1 e^{\lambda_1 t}$, con fórmulas similares para $y_2, \ldots, y_n$. Entonces

$$\mathbf{y}(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}, \quad \text{donde} \quad \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{y}(0) = P^{-1}\mathbf{x}(0) = P^{-1}\mathbf{x}_0$$

Para obtener la solución general $\mathbf{x}$ del sistema original, calcule

$$\begin{aligned} \mathbf{x}(t) &= P\mathbf{y}(t) = [\,\mathbf{v}_1 \ \cdots \ \mathbf{v}_n\,]\,\mathbf{y}(t) \\ &= c_1\mathbf{v}_1 e^{\lambda_1 t} + \cdots + c_n\mathbf{v}_n e^{\lambda_n t} \end{aligned}$$

Ésta es la expansión de la función propia estructurada como en el ejemplo 1.

Valores propios complejos

En el ejemplo siguiente, una matriz real A tiene un par de valores propios complejos λ y $\overline{\lambda}$, con los vectores propios complejos asociados $\mathbf{v}$ y $\overline{\mathbf{v}}$. (De la sección 5.5, recuerde que para una matriz real, los valores propios complejos y los vectores propios asociados se presentan en pares conjugados.) Así que dos soluciones de $\mathbf{x}' = A\mathbf{x}$ son

$$\mathbf{x}_1(t) = \mathbf{v}e^{\lambda t} \quad \text{y} \quad \mathbf{x}_2(t) = \overline{\mathbf{v}}e^{\overline{\lambda} t} \tag{6}$$

Puede mostrarse que $\mathbf{x}_2(t) = \overline{\mathbf{x}_1(t)}$ usando una representación en serie de potencias para la función exponencial compleja. Aunque las funciones propias complejas $\mathbf{x}_1$ y $\mathbf{x}_2$ son convenientes para realizar algunos cálculos (sobre todo en ingeniería eléctrica), para muchos propósitos resultan más apropiadas las funciones reales. Por fortuna, las partes real e imaginaria de $\mathbf{x}_1$ son soluciones (reales) de $\mathbf{x}' = A\mathbf{x}$, porque son combinaciones lineales de las soluciones de (6):

$$\text{Re}(\mathbf{v}e^{\lambda t}) = \frac{1}{2}\left[\,\mathbf{x}_1(t) + \overline{\mathbf{x}_1(t)}\,\right], \qquad \text{Im}(\mathbf{v}e^{\lambda t}) = \frac{1}{2i}\left[\,\mathbf{x}_1(t) - \overline{\mathbf{x}_1(t)}\,\right]$$

Para entender la naturaleza de $\text{Re}(\mathbf{v}e^{\lambda t})$, recuerde de sus lecciones de cálculo que para cualquier número x se puede encontrar la función exponencial e^x a partir de la serie de potencias:

$$e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots$$

Es posible usar esta serie para definir $e^{\lambda t}$ cuando λ es complejo:

$$e^{\lambda t} = 1 + (\lambda t) + \frac{1}{2!}(\lambda t)^2 + \cdots + \frac{1}{n!}(\lambda t)^n + \cdots$$

Si se escribe $\lambda = a + bi$ (con a y b reales), y se usan series de potencias semejantes para las funciones seno y coseno, se puede mostrar que

$$e^{(a+bi)t} = e^{at} \cdot e^{ibt} = e^{at}(\cos bt + i \operatorname{sen} bt) \tag{7}$$

Por lo tanto,

$$\begin{aligned} \mathbf{v}e^{\lambda t} &= (\text{Re}\,\mathbf{v} + i\,\text{Im}\,\mathbf{v}) \cdot e^{at}(\cos bt + i \operatorname{sen} bt) \\ &= [\,(\text{Re}\,\mathbf{v})\cos bt - (\text{Im}\,\mathbf{v})\operatorname{sen} bt\,]\,e^{at} \\ &\quad + i\,[\,(\text{Re}\,\mathbf{v})\operatorname{sen} bt + (\text{Im}\,\mathbf{v})\cos bt\,]\,e^{at} \end{aligned}$$

Así, dos soluciones reales de $\mathbf{x}' = A\mathbf{x}$ son

$$\boxed{\begin{aligned} \mathbf{y}_1(t) &= \text{Re}\,\mathbf{x}_1(t) = [\,(\text{Re}\,\mathbf{v})\cos bt - (\text{Im}\,\mathbf{v})\operatorname{sen} bt\,]\,e^{at} \\ \mathbf{y}_2(t) &= \text{Im}\,\mathbf{x}_1(t) = [\,(\text{Re}\,\mathbf{v})\operatorname{sen} bt + (\text{Im}\,\mathbf{v})\cos bt\,]\,e^{at} \end{aligned}}$$

Puede mostrarse que $\mathbf{y}_1$ y $\mathbf{y}_2$ son funciones linealmente independientes (cuando $b \neq 0$).[1]

[1]Como $\mathbf{x}_2(t)$ es el conjugado complejo de $\mathbf{x}_1(t)$, las partes real e imaginaria de $\mathbf{x}_2(t)$ son $\mathbf{y}_1(t)$ y $-\mathbf{y}_2(t)$, respectivamente. Entonces se puede usar sea $\mathbf{x}_1(t)$ o $\mathbf{x}_2(t)$, pero no ambos, para producir dos soluciones reales linealmente independientes de $\mathbf{x}' = A\mathbf{x}$.

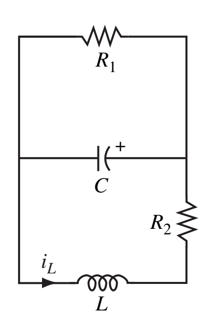


FIGURA 4

EJEMPLO 3 El circuito de la figura 4 puede describirse mediante la ecuación

$$\begin{bmatrix} i'_L \\ v'_C \end{bmatrix} = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1C) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

donde i_L es la corriente que pasa por el inductor L y v_C es la caída de voltaje a través del condensador C. Suponga que R_1 es de 5 ohms, R_2 de .8 ohms, C de .1 farad, y L de .4 henrys. Encuentre fórmulas para i_L y v_C si la corriente inicial a través del inductor es de 3 amperes y el voltaje inicial en el capacitor es de 3 volts.

Solución Para los datos proporcionados, $\begin{bmatrix} -2 & -2.5 \\ 10 & -2 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. El método de la sección 5.5 produce el valor propio $\lambda = -2 + 5i$ y el correspondiente vector propio $\mathbf{v}_1 = \begin{bmatrix} i \\ 2 \end{bmatrix}$. Las soluciones complejas de $\mathbf{x}' = A\mathbf{x}$ son combinaciones lineales de

$$\mathbf{x}_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{(-2+5i)t} \quad \text{y} \quad \mathbf{x}_2(t) = \begin{bmatrix} -i \\ 2 \end{bmatrix} e^{(-2-5i)t}$$

Enseguida, use (7) y escriba

$$\mathbf{x}_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{-2t}(\cos 5t + i \operatorname{sen} 5t)$$

Las partes real e imaginaria de $\mathbf{x}_1$ proporcionan soluciones reales:

$$\mathbf{y}_1(t) = \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t}, \qquad \mathbf{y}_2(t) = \begin{bmatrix} \cos 5t \\ 2\operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

Como $\mathbf{y}_1$ y $\mathbf{y}_2$ son funciones linealmente independientes, forman una base para el espacio vectorial real bidimensional de soluciones de $\mathbf{x}' = A\mathbf{x}$. Entonces la solución general es

$$\mathbf{x}(t) = c_1 \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \cos 5t \\ 2\operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

Para satisfacer $\mathbf{x}(0) = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, se necesita $c_1 \begin{bmatrix} 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, lo cual conduce a $c_1 = 1.5$ y $c_2 = 3$. Así que

$$\mathbf{x}(t) = 1.5 \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + 3 \begin{bmatrix} \cos 5t \\ 2\operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

o bien

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -1.5\operatorname{sen} 5t + 3\cos 5t \\ 3\cos 5t + 6\operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

Vea la figura 5.

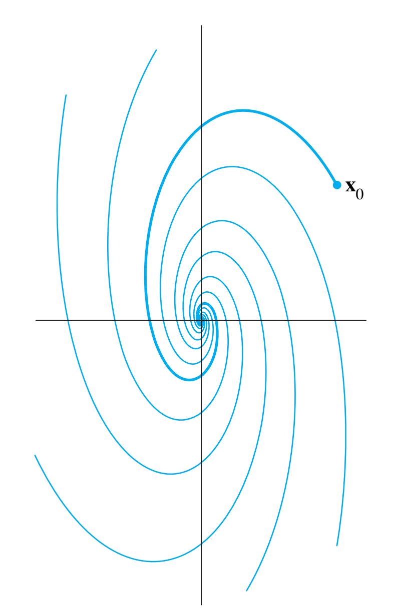


FIGURA 5
El origen como punto espiral.

En la figura 5, el origen se llama **punto espiral** del sistema dinámico. La rotación es causada por las funciones seno y coseno que surgen de un valor propio complejo. Las trayectorias describen una espiral hacia dentro porque el factor e^{-2t} tiende a cero. Recuerde que -2 es la parte real del valor propio analizado en el ejemplo 3. Cuando

A tiene un valor propio complejo con la parte real positiva, las trayectorias describen una espiral hacia fuera. Si la parte real del valor propio es cero, las trayectorias forman elipses alrededor del origen.

PROBLEMAS DE PRÁCTICA

Una matriz A real de 3×3 tiene valores propios $-.5$, $.2 + .3i$, y $.2 - .3i$, con los correspondientes vectores propios

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix}, \qquad \text{y} \qquad \mathbf{v}_3 = \begin{bmatrix} 1-2i \\ -4i \\ 2 \end{bmatrix}$$

1. ¿Puede diagonalizarse A como $A = PDP^{-1}$ utilizando matrices complejas?
2. Escriba la solución general de $\mathbf{x}' = A\mathbf{x}$ usando funciones propias complejas, y luego encuentre la solución general real.
3. Describa las formas de trayectorias típicas.

5.7 EJERCICIOS

1. Una partícula que se mueve en un campo de fuerza plano tiene un vector de posición $\mathbf{x}$ que satisface $\mathbf{x}' = A\mathbf{x}$. La matriz A de 2×2 tiene valores propios 4 y 2, con vectores propios correspondientes $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Encuentre la posición de la partícula en el tiempo t, suponiendo que $\mathbf{x}(0) = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$.

2. Sea A una matriz de 2×2 con valores propios -3 y -1 y los correspondientes vectores propios $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Sea $\mathbf{x}(t)$ la posición de una partícula en el tiempo t. Resuelva el problema de valor inicial $\mathbf{x}' = A\mathbf{x}$, $\mathbf{x}(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

En los ejercicios 3 a 6, resuelva el problema de valor inicial $\mathbf{x}'(t) = A\mathbf{x}(t)$ para $t \geq 0$, con $\mathbf{x}(0) = (3, 2)$. Clasifique la naturaleza del origen como un atractor, repulsor, o punto silla del sistema dinámico descrito mediante $\mathbf{x}' = A\mathbf{x}$. Encuentre los sentidos de mayor atracción y/o repulsión. Cuando el origen sea un punto silla, trace las trayectorias características.

3. $A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$
4. $A = \begin{bmatrix} -2 & -5 \\ 1 & 4 \end{bmatrix}$
5. $A = \begin{bmatrix} 7 & -1 \\ 3 & 3 \end{bmatrix}$
6. $A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}$

En los ejercicios 7 y 8, realice un cambio de variable que desacople la ecuación $\mathbf{x}' = A\mathbf{x}$. Escriba la ecuación $\mathbf{x}(t) = P\mathbf{y}(t)$ y muestre los cálculos que conducen al sistema desacoplado $\mathbf{y}' = D\mathbf{y}$, especificando P y D.

7. A como en el ejercicio 5.
8. A como en el ejercicio 6.

En los ejercicios 9 a 18, estructure la solución general de $\mathbf{x}' = A\mathbf{x}$ usando funciones propias complejas, y luego obtenga la solución general real. Describa las formas de trayectorias características.

9. $A = \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix}$
10. $A = \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$
11. $A = \begin{bmatrix} -3 & -9 \\ 2 & 3 \end{bmatrix}$
12. $A = \begin{bmatrix} -7 & 10 \\ -4 & 5 \end{bmatrix}$
13. $A = \begin{bmatrix} 4 & -3 \\ 6 & -2 \end{bmatrix}$
14. $A = \begin{bmatrix} -2 & 1 \\ -8 & 2 \end{bmatrix}$
15. **[M]** $A = \begin{bmatrix} -8 & -12 & -6 \\ 2 & 1 & 2 \\ 7 & 12 & 5 \end{bmatrix}$
16. **[M]** $A = \begin{bmatrix} -6 & -11 & 16 \\ 2 & 5 & -4 \\ -4 & -5 & 10 \end{bmatrix}$
17. **[M]** $A = \begin{bmatrix} 30 & 64 & 23 \\ -11 & -23 & -9 \\ 6 & 15 & 4 \end{bmatrix}$

18. **[M]** $A = \begin{bmatrix} 53 & -30 & -2 \\ 90 & -52 & -3 \\ 20 & -10 & 2 \end{bmatrix}$

19. **[M]** Encuentre las fórmulas para los voltajes v_1 y v_2 (como funciones del tiempo t) para el circuito del ejemplo 1, suponiendo que $R_1 = 1/5$ ohms, $R_2 = 1/3$ ohms, $C_1 = 4$ farads, $C_2 = 3$ farads, y que la carga inicial en cada capacitor es de 4 volts.

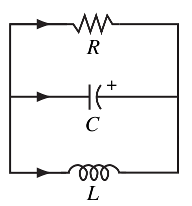
20. **[M]** Encuentre las fórmulas para los voltajes v_1 y v_2 para el circuito del ejemplo 1, suponiendo que $R_1 = 1/15$ ohms, $R_2 = 1/3$ ohms, $C_1 = 9$ farads, $C_2 = 2$ farads, y que la carga inicial en cada capacitor es de 3 volts.

21. **[M]** Encuentre las fórmulas para la corriente i_L y el voltaje v_C para el circuito del ejemplo 3, suponiendo que $R_1 = 1$ ohm, $R_2 = .125$ ohms, $C = .2$ farads, $L = .125$ henrys, la corriente inicial es de 0 amperes y el voltaje inicial de 15 volts.

22. **[M]** El circuito de la figura 6 se describe mediante la ecuación

$$\begin{bmatrix} i'_L \\ v'_C \end{bmatrix} = \begin{bmatrix} 0 & 1/L \\ -1/C & -1/(RC) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

donde i_L es la corriente que pasa por el inductor L y v_C es la caída de voltaje a través del condensador C. Encuentre las fórmulas para i_L y v_C cuando $R = .5$ ohms, $C = 2.5$ farads, $L = .5$ henrys, la corriente inicial es de 0 amperes y el voltaje inicial de 12 volts.



SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sí, la matriz de 3×3 es diagonalizable porque tiene tres valores propios distintos. El teorema 2 de la sección 5.1 y el teorema 5 de la sección 5.3 son válidos cuando se usan escalares complejos. (Las demostraciones son esencialmente las mismas que para los escalares reales.)

2. La solución general tiene la forma

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix} e^{(.2+.3i)t} + c_3 \begin{bmatrix} 1-2i \\ -4i \\ 2 \end{bmatrix} e^{(.2-.3i)t}$$

Aquí los escalares c_1, c_2, c_3 pueden ser cualesquiera números complejos. El primer término de $\mathbf{x}(t)$ es real. Pueden producirse dos soluciones reales más usando las partes real e imaginaria del segundo término de $\mathbf{x}(t)$:

$$\begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix} e^{.2t}(\cos .3t + i \operatorname{sen} .3t)$$

La solución real general tiene la siguiente forma, con escalares *reales* c_1, c_2 y c_3:

$$c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} \cos .3t - 2 \operatorname{sen} .3t \\ -4 \operatorname{sen} .3t \\ 2 \cos .3t \end{bmatrix} e^{.2t} + c_3 \begin{bmatrix} \operatorname{sen} .3t + 2 \cos .3t \\ 4 \cos .3t \\ 2 \operatorname{sen} .3t \end{bmatrix} e^{.2t}$$

3. Cualquier solución con $c_2 = c_3 = 0$ es atraída hacia el origen a causa del factor exponencial negativo. Otras soluciones tienen componentes que crecen ilimitadamente, y las trayectorias describen una espiral hacia fuera.

Se recomienda tener cuidado de no confundir este problema con uno de la sección 5.6. Allí la condición para la atracción hacia **0** era que un valor propio tuviera magnitud menor que 1, para hacer $|\lambda|^k \to 0$. Aquí la parte real del valor propio debe ser negativa, para hacer $e^{\lambda t} \to 0$.

5.8 ESTIMACIONES ITERATIVAS PARA VALORES PROPIOS

En las aplicaciones científicas del álgebra lineal, rara vez se conocen los valores propios con precisión. Por fortuna, una aproximación numérica cercana casi siempre resulta satisfactoria. De hecho, algunas aplicaciones requieren sólo de una aproximación burda al valor propio más grande. El primer algoritmo descrito a continuación puede funcionar bien para este caso; asimismo, proporciona la base para un método más potente que también puede entregar estimaciones rápidas de otros valores propios.

El método de potencias

El método de potencias se aplica a una matriz A de $n \times n$ con un **valor propio estrictamente dominante** λ_1, lo cual significa que λ_1 debe ser mayor en valor absoluto que cualquier otro valor propio. En este caso, el método de potencias produce una sucesión escalar que se aproxima a λ_1 y una sucesión vectorial que se aproxima al correspondiente vector propio. Los antecedentes de este método se basan en la descomposición del vector propio usada al principio de la sección 5.6.

En aras de la simplicidad, suponga que A es diagonalizable y que $\mathbb{R}^n$ tiene una base de vectores propios $\mathbf{v}_1, \ldots, \mathbf{v}_n$, acomodados de manera que sus valores propios correspondientes $\lambda_1, \ldots, \lambda_n$ disminuyan de tamaño, con el valor propio estrictamente dominante en primer lugar. Esto es,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \tag{1}$$

Estrictamente mayor

Como se vio en la ecuación (2) de la sección 5.6, si $\mathbf{x}$ en $\mathbb{R}^n$ se escribe como $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, entonces

$$A^k\mathbf{x} = c_1(\lambda_1)^k\mathbf{v}_1 + c_2(\lambda_2)^k\mathbf{v}_2 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n \quad (k = 1, 2, \ldots)$$

Suponga que $c_1 \neq 0$. Entonces, al dividir entre $(\lambda_1)^k$,

$$\frac{1}{(\lambda_1)^k}A^k\mathbf{x} = c_1\mathbf{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v}_n \quad (k = 1, 2, \ldots) \tag{2}$$

A partir de (1), todas las fracciones $\lambda_2/\lambda_1, \ldots, \lambda_n/\lambda_1$ son de magnitud menor que 1, de manera que sus potencias van al cero. Por lo tanto,

$$(\lambda_1)^{-k}A^k\mathbf{x} \to c_1\mathbf{v}_1 \quad \text{as } k \to \infty \tag{3}$$

Entonces, para k grande, un múltiplo escalar de $A^k\mathbf{x}$ determina casi la misma *dirección* que el vector propio $c_1\mathbf{v}_1$. Puesto que los múltiplos escalares positivos no cambian el sentido de un vector, $A^k\mathbf{x}$ apunta casi en la misma dirección que $\mathbf{v}_1$ o $-\mathbf{v}_1$, dado que $c_1 \neq 0$.

EJEMPLO 1 Sean $A = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, y $\mathbf{x} = \begin{bmatrix} -.5 \\ 1 \end{bmatrix}$. Entonces A tiene valores propios 2 y 1, y el espacio propio para $\lambda_1 = 2$ es la línea que pasa por $\mathbf{0}$ y $\mathbf{v}_1$.

Para $k = 0, \ldots, 8$, calcule $A^k\mathbf{x}$ y construya la línea que pasa por $\mathbf{0}$ y $A^k\mathbf{x}$. ¿Qué ocurre al aumentar k?

Solución Los primeros tres cálculos son

$$A\mathbf{x} = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} -.5 \\ 1 \end{bmatrix} = \begin{bmatrix} -.1 \\ 1.1 \end{bmatrix}$$

$$A^2\mathbf{x} = A(A\mathbf{x}) = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} -.1 \\ 1.1 \end{bmatrix} = \begin{bmatrix} .7 \\ 1.3 \end{bmatrix}$$

$$A^3\mathbf{x} = A(A^2\mathbf{x}) = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} .7 \\ 1.3 \end{bmatrix} = \begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$$

La tabla 1 puede completarse con cálculos análogos.

TABLA 1 Iteraciones de un vector

k	0	1	2	3	4	5	6	7	8
$A^k\mathbf{x}$	$\begin{bmatrix} -.5 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -.1 \\ 1.1 \end{bmatrix}$	$\begin{bmatrix} .7 \\ 1.3 \end{bmatrix}$	$\begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$	$\begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$	$\begin{bmatrix} 11.9 \\ 4.1 \end{bmatrix}$	$\begin{bmatrix} 24.7 \\ 7.3 \end{bmatrix}$	$\begin{bmatrix} 50.3 \\ 13.7 \end{bmatrix}$	$\begin{bmatrix} 101.5 \\ 26.5 \end{bmatrix}$

Los vectores $\mathbf{x}, A\mathbf{x}, \ldots, A^4\mathbf{x}$ se muestran en la figura 1. Los otros vectores se vuelven demasiado largos como para exhibirlos. No obstante, se han trazado segmentos de línea que muestran las direcciones de esos vectores. De hecho, lo que realmente se desea observar es el sentido de los vectores, no los vectores mismos. Las líneas parecen estar aproximándose a la línea que representa el espacio propio generado por $\mathbf{v}_1$. Con mayor precisión, el ángulo entre la línea (subespacio) determinada por $A^k\mathbf{x}$ y la línea (espacio propio) determinada por $\mathbf{v}_1$ tiende a cero conforme $k \to \infty$.

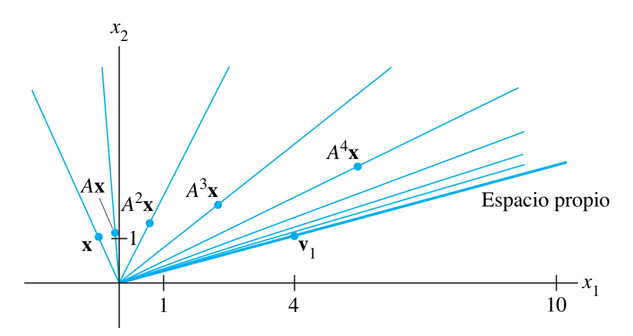


FIGURA 1 Direcciones determinadas por $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \ldots, A^7\mathbf{x}$.

Los vectores $(\lambda_1)^{-k}A^k\mathbf{x}$ de (3) están escalados para hacerlos converger hacia $c_1\mathbf{v}_1$, a condición de que $c_1 \neq 0$. No es posible escalar de esta manera $A^k\mathbf{x}$ porque no se conoce λ_1. Pero puede escalarse cada $A^k\mathbf{x}$ para hacer que su entrada mayor sea 1. Resulta que la sucesión $\{\mathbf{x}_k\}$ que se obtiene convergerá hacia un múltiplo de $\mathbf{v}_1$ cuya entrada mayor sea 1. La figura 2 muestra la sucesión escalada para el ejemplo 1. El valor propio λ_1 también puede estimarse a partir de la sucesión $\{\mathbf{x}_k\}$. Cuando $\mathbf{x}_k$ es cercano a un vector propio para λ_1, el vector $A\mathbf{x}_k$ es cercano a $\lambda_1\mathbf{x}_k$ siendo cada entrada de $A\mathbf{x}_k$ aproximadamente λ_1 veces la entrada correspondiente de $\mathbf{x}_k$. Como la entrada mayor de $\mathbf{x}_k$ es 1, la

entrada mayor de $A\mathbf{x}_k$ es cercana a λ_1. (Se omiten las demostraciones detalladas de estos enunciados.)

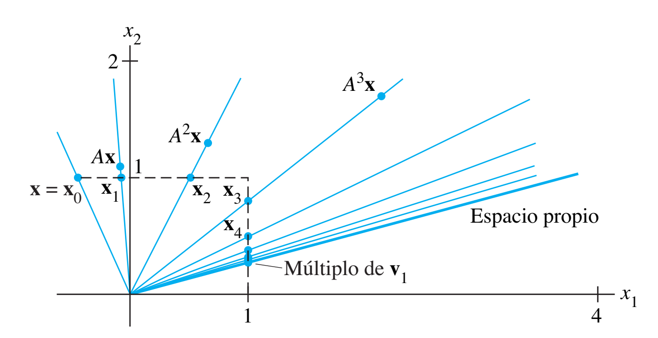


FIGURA 2 Múltiplos escalados de $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \ldots, A^7\mathbf{x}$.

EL MÉTODO DE POTENCIAS PARA LA ESTIMACIÓN DE UN VALOR PROPIO ESTRICTAMENTE DOMINANTE

1. Seleccione un vector inicial $\mathbf{x}_0$ cuya entrada mayor sea 1.
2. Para $k = 0, 1, \ldots,$
   a. Calcule $A\mathbf{x}_k$.
   b. Sea μ_k una entrada de $A\mathbf{x}_k$ cuyo valor absoluto sea lo más grande posible.
   c. Calcule $\mathbf{x}_{k+1} = (1/\mu_k)A\mathbf{x}_k$.
3. Para casi todas las elecciones de $\mathbf{x}_0$, la sucesión $\{\mu_k\}$ se aproxima al valor propio dominante y la sucesión $\{\mathbf{x}_k\}$ se aproxima al correspondiente vector propio.

EJEMPLO 2 Aplique el método de potencias a $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ con $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Debe aplicarse el método hasta $k = 5$ y estimar el valor propio dominante y un correspondiente vector propio de A.

Solución Los cálculos de este ejemplo y del siguiente se hicieron con MATLAB, que calcula con una precisión de 16 dígitos, aunque aquí se muestran sólo unas cuantas cifras significativas. Para comenzar, calcule $A\mathbf{x}_0$ e identifique la entrada más grande μ_0 en $A\mathbf{x}_0$:

$$A\mathbf{x}_0 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \quad \mu_0 = 5$$

Escale $A\mathbf{x}_0$ mediante $1/\mu_0$ para obtener $\mathbf{x}_1$, calcule $A\mathbf{x}$, e identifique la máxima entrada en $A\mathbf{x}_1$:

$$\mathbf{x}_1 = \frac{1}{\mu_0}A\mathbf{x}_0 = \frac{1}{5}\begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ .4 \end{bmatrix}$$

$$A\mathbf{x}_1 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1.8 \end{bmatrix}, \quad \mu_1 = 8$$

Escale $A\mathbf{x}_1$ mediante $1/\mu_1$ para obtener $\mathbf{x}_2$, calcule $A\mathbf{x}_2$ e identifique la entrada máxima en $A\mathbf{x}_2$:

$$\mathbf{x}_2 = \frac{1}{\mu_1}A\mathbf{x}_1 = \frac{1}{8}\begin{bmatrix} 8 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1 \\ .225 \end{bmatrix}$$

$$A\mathbf{x}_2 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .225 \end{bmatrix} = \begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}, \quad \mu_2 = 7.125$$

Escale $A\mathbf{x}_2$ mediante $1/\mu_2$ para obtener $\mathbf{x}_3$, y así sucesivamente. Los resultados de los cálculos efectuados en MATLAB para las primeras cinco iteraciones están acomodados como se muestra en la tabla 2.

TABLA 2 El método de potencias para el ejemplo 2

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .4 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .225 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2035 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2005 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .20007 \end{bmatrix}$
$A\mathbf{x}_k$	$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 8 \\ 1.8 \end{bmatrix}$	$\begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}$	$\begin{bmatrix} 7.0175 \\ 1.4070 \end{bmatrix}$	$\begin{bmatrix} 7.0025 \\ 1.4010 \end{bmatrix}$	$\begin{bmatrix} 7.00036 \\ 1.40014 \end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

La evidencia expuesta en la tabla 2 claramente sugiere que $\{\mathbf{x}_k\}$ se aproxima a (1, .2) y que $\{\mu_k\}$ se aproxima a 7. Si esto es así, entonces (1, .2) es un vector propio y 7 es el valor propio dominante. Lo cual se verifica fácilmente al calcular

$$A\begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 7 \\ 1.4 \end{bmatrix} = 7\begin{bmatrix} 1 \\ .2 \end{bmatrix}$$

La sucesión $\{\mu_k\}$ del ejemplo 2 converge rápidamente a $\lambda_1 = 7$ porque el segundo valor propio de A era mucho menor. (De hecho, $\lambda_2 = 1$.) En general, la tasa de convergencia depende de la razón $|\lambda_2/\lambda_1|$, porque el vector $c_2(\lambda_2/\lambda_1)^k\mathbf{v}_2$ en (2) es la principal fuente de error cuando se usa una versión escalada de $A^k\mathbf{x}$ como una estimación de $c_1\mathbf{v}_1$. (Es probable que las otras fracciones λ_j/λ_1 sean menores.) Si $|\lambda_2/\lambda_1|$ es cercana a 1, entonces $\{\mu_k\}$ y $\{\mathbf{x}_k\}$ pueden converger muy lentamente y podrían preferirse otros métodos de aproximación.

Con el método de potencias hay una ligera posibilidad de que una selección aleatoria del vector inicial $\mathbf{x}$ no tenga componente en la dirección $\mathbf{v}_1$ (cuando $c_1 = 0$). Pero es probable que los errores de redondeo de la computadora durante los cálculos de las $\mathbf{x}_k$ generen un vector con, por lo menos, un pequeño componente en la dirección de $\mathbf{v}_1$. Si eso pasa, las $\mathbf{x}_k$ empezarán a converger a un múltiplo de $\mathbf{v}_1$.

El método de potencias inversas

Este método proporciona una aproximación para *cualquier* valor propio, siempre que se conozca una buena estimación inicial α del valor propio λ. En este caso, se supone que

$B = (A - \alpha I)^{-1}$ y se aplica el método de potencias a B. Se puede demostrar que si los valores propios de A son $\lambda_1, \ldots, \lambda_n$, entonces los valores propios de B son

$$\frac{1}{\lambda_1 - \alpha}, \quad \frac{1}{\lambda_2 - \alpha}, \quad \ldots, \quad \frac{1}{\lambda_n - \alpha}$$

y los correspondientes vectores propios son los mismos que los de A. (Vea los ejercicios 15 y 16.)

Suponga, por ejemplo, que α es más cercana a λ_2 que a los otros valores propios de A. Entonces $1/(\lambda_2 - \alpha)$ será un valor propio estrictamente dominante de B. Si α es muy cercana a λ_2, entonces $1/(\lambda_2 - \alpha)$ es *mucho* mayor que los otros valores propios de B, y el método de potencias inversas produce una aproximación muy rápida a λ_2 para casi cualquier selección de $\mathbf{x}_0$. El siguiente algoritmo proporciona los detalles.

EL MÉTODO DE POTENCIAS INVERSAS PARA ESTIMAR UN VALOR PROPIO λ DE A

1. Seleccione una α estimada inicial lo suficientemente cercana a λ.

2. Seleccione un vector inicial $\mathbf{x}_0$ cuya entrada más grande sea 1.

3. Para $k = 0, 1, \ldots,$

   a. Resuelva $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$ para $\mathbf{y}_k$.

   b. Sea μ_k una entrada en $\mathbf{y}_k$ con valor absoluto tan grande como sea posible.

   c. Calcule $\nu_k = \alpha + (1/\mu_k)$.

   d. Calcule $\mathbf{x}_{k+1} = (1/\mu_k)\mathbf{y}_k$.

4. Para casi todas las elecciones de $\mathbf{x}_0$, la sucesión $\{\nu_k\}$ se aproxima al valor propio λ de A, y la sucesión $\{\mathbf{x}_k\}$ se aproxima a un vector propio correspondiente.

Observe que B, o más bien $(A - \alpha I)^{-1}$, no aparece en el algoritmo. En lugar de calcular $(A - \alpha I)^{-1}\mathbf{x}_k$ para obtener el siguiente vector de la sucesión, es mejor *resolver* la ecuación $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$ para $\mathbf{y}_k$ (y luego escalar $\mathbf{y}_k$ para producir $\mathbf{x}_{k+1}$). Puesto que debe resolverse esta ecuación de $\mathbf{y}_k$ para cada k, una factorización LU de $A - \alpha I$ acelerará el proceso.

EJEMPLO 3 No es raro que en algunas aplicaciones sea necesario conocer el valor propio más pequeño de una matriz A y tener a la mano estimaciones burdas de los valores propios. Suponga que 21, 3.3, y 1.9 son estimaciones de los valores propios de la matriz A siguiente. Encuentre el valor propio más pequeño, precisando hasta seis lugares decimales.

$$A = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 4 \\ -4 & 5 & 4 \end{bmatrix}$$

Solución Los dos valores propios más pequeños parecen estar muy cerca uno del otro, así que se utiliza el método de potencias inversas para $A - 1.9I$. En la tabla 3 se muestran los resultados de un cálculo efectuado en MATLAB. Aquí $\mathbf{x}_0$ fue seleccionado arbitrariamente, $\mathbf{y}_k = (A - 1.9I)^{-1}\mathbf{x}_k$, μ_k es la entrada más grande en $\mathbf{y}_k$, $\nu_k = 1.9 + 1/\mu_k$, y $\mathbf{x}_{k+1} = (1/\mu_k)\mathbf{y}_k$. Resulta que la estimación del valor propio inicial fue bastante buena, y la sucesión de potencias inversas convergió muy rápido. El valor propio más pequeño es exactamente 2.

TABLA 3 El método de potencias inversas

k	0	1	2	3	4
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5736 \\ .0646 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5054 \\ .0045 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5004 \\ .0003 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .50003 \\ .00002 \\ 1 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 4.45 \\ .50 \\ 7.76 \end{bmatrix}$	$\begin{bmatrix} 5.0131 \\ .0442 \\ 9.9197 \end{bmatrix}$	$\begin{bmatrix} 5.0012 \\ .0031 \\ 9.9949 \end{bmatrix}$	$\begin{bmatrix} 5.0001 \\ .0002 \\ 9.9996 \end{bmatrix}$	$\begin{bmatrix} 5.000006 \\ .000015 \\ 9.999975 \end{bmatrix}$
μ_k	7.76	9.9197	9.9949	9.9996	9.999975
ν_k	2.03	2.0008	2.00005	2.000004	2.0000002

Si no se dispone de una estimación para el valor propio más pequeño de una matriz, es posible tomar simplemente $\alpha = 0$ en el método de potencias inversas. Esta elección de α funciona razonablemente bien si el valor propio más pequeño está mucho más cerca de cero que de los otros valores propios.

Los dos algoritmos presentados en esta sección son herramientas prácticas para emplear en muchas situaciones sencillas, y ofrecen una introducción al problema de la estimación del valor propio. Un método iterativo más eficaz y ampliamente usado es el algoritmo QR. Por ejemplo, este algoritmo es el corazón de la orden `eig(A)` de MATLAB, el cual calcula rápidamente los vectores y valores propios de A. Una breve descripción del algoritmo QR fue proporcionada en los ejercicios de la sección 5.2. Se pueden encontrar mayores detalles en casi todos los textos modernos de análisis numérico.

PROBLEMA DE PRÁCTICA

¿Cómo puede saberse si un vector $\mathbf{x}$ dado es una buena aproximación a un vector propio de una matriz A? Y si lo es, ¿cómo se estimaría el correspondiente valor propio? Experimente con

$$A = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} 1.0 \\ -4.3 \\ 8.1 \end{bmatrix}$$

5.8 Ejercicios

En los ejercicios 1 a 4, a la matriz A le sigue una sucesión $\{\mathbf{x}_k\}$ producida mediante el método de potencias. Use estos datos para estimar el mayor valor propio de A, y proporcione un vector propio correspondiente.

1. $A = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$; $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .25 \end{bmatrix}, \begin{bmatrix} 1 \\ .3158 \end{bmatrix}, \begin{bmatrix} 1 \\ .3298 \end{bmatrix}, \begin{bmatrix} 1 \\ .3326 \end{bmatrix}$

2. $A = \begin{bmatrix} 1.8 & -.8 \\ -3.2 & 4.2 \end{bmatrix}$;

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -.5625 \\ 1 \end{bmatrix}, \begin{bmatrix} -.3021 \\ 1 \end{bmatrix}, \begin{bmatrix} -.2601 \\ 1 \end{bmatrix}, \begin{bmatrix} -.2520 \\ 1 \end{bmatrix}$

3. $A = \begin{bmatrix} .5 & .2 \\ .4 & .7 \end{bmatrix}$;

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .8 \end{bmatrix}, \begin{bmatrix} .6875 \\ 1 \end{bmatrix}, \begin{bmatrix} .5577 \\ 1 \end{bmatrix}, \begin{bmatrix} .5188 \\ 1 \end{bmatrix}$

4. $A = \begin{bmatrix} 4.1 & -6 \\ 3 & -4.4 \end{bmatrix}$;

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ .7368 \end{bmatrix}, \begin{bmatrix} 1 \\ .7541 \end{bmatrix}, \begin{bmatrix} 1 \\ .7490 \end{bmatrix}, \begin{bmatrix} 1 \\ .7502 \end{bmatrix}$

5. Sea $A = \begin{bmatrix} 15 & 16 \\ -20 & -21 \end{bmatrix}$. Los vectores $\mathbf{x}, \ldots, A^5\mathbf{x}$ son $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 31 \\ -41 \end{bmatrix}, \begin{bmatrix} -191 \\ 241 \end{bmatrix}, \begin{bmatrix} 991 \\ -1241 \end{bmatrix}, \begin{bmatrix} -4991 \\ 6241 \end{bmatrix}, \begin{bmatrix} 24991 \\ -31241 \end{bmatrix}$. Encuentre un vector con un 1 en la segunda entrada que sea un vector propio de A. Use cuatro posiciones decimales. Compruebe la estimación, y proporcione una estimación para el valor propio dominante de A.

6. Sea $A = \begin{bmatrix} -2 & -3 \\ 6 & 7 \end{bmatrix}$. Repita el ejercicio 5 usando la siguiente sucesión $\mathbf{x}, A\mathbf{x}, \ldots, A^5\mathbf{x}$.

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 13 \end{bmatrix}, \begin{bmatrix} -29 \\ 61 \end{bmatrix}, \begin{bmatrix} -125 \\ 253 \end{bmatrix}, \begin{bmatrix} -509 \\ 1021 \end{bmatrix}, \begin{bmatrix} -2045 \\ 4093 \end{bmatrix}$

[M] En los ejercicios 7 a 12 se requiere MATLAB u otra ayuda de computadora. En los ejercicios 7 y 8, utilice el método de potencias con el $\mathbf{x}_0$ dado. Enliste $\{\mathbf{x}_k\}$ y $\{\mu_k\}$ para $k = 1, \ldots, 5$. En los ejercicios 9 y 10, enliste μ_5 y μ_6.

7. $A = \begin{bmatrix} 6 & 7 \\ 8 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

8. $A = \begin{bmatrix} 2 & 1 \\ 4 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

9. $A = \begin{bmatrix} 8 & 0 & 12 \\ 1 & -2 & 1 \\ 0 & 3 & 0 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 9 \\ 0 & 1 & 9 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

Es posible efectuar otra estimación de un valor propio cuando se dispone de un vector propio aproximado. Observe que si $A\mathbf{x} = \lambda\mathbf{x}$, entonces $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T(\lambda\mathbf{x}) = \lambda(\mathbf{x}^T\mathbf{x})$, y el **cociente de Rayleigh**

$$R(\mathbf{x}) = \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

es igual a λ. Si $\mathbf{x}$ es cercano a un vector propio para λ, entonces este cociente es cercano a λ. Cuando A sea una matriz simétrica ($A^T = A$), el cociente de Rayleigh $R(\mathbf{x}_k) = (\mathbf{x}_k^T A\mathbf{x}_k)/(\mathbf{x}_k^T\mathbf{x}_k)$ tendrá una precisión de aproximadamente el doble de dígitos que el factor de escalamiento μ_k en el método de potencias. Verifique el aumento de precisión en los ejercicios 11 y 12 calculando μ_k y $R(\mathbf{x}_k)$ para $k = 1, \ldots, 4$.

11. $A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

12. $A = \begin{bmatrix} -3 & 2 \\ 2 & 0 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Los ejercicios 13 y 14 se aplican a una matriz A de 3×3 cuyos valores propios se estiman en 4, -4, y 3.

13. Si los valores propios cercanos a 4 y -4 tienen diferentes valores absolutos, ¿funcionará el método de potencias? ¿Es probable que resulte útil?

14. Suponga que los valores propios cercanos a 4 y -4 tienen exactamente el mismo valor absoluto. Describa cómo podría obtenerse una sucesión que estime el valor propio cercano a 4.

15. Suponga $A\mathbf{x} = \lambda\mathbf{x}$ con $\mathbf{x} \neq \mathbf{0}$. Sea α un escalar diferente de los valores propios de A, y sea $B = (A - \alpha I)^{-1}$. Reste $\alpha\mathbf{x}$ de ambos miembros de la ecuación $A\mathbf{x} = \lambda\mathbf{x}$, y utilice álgebra para mostrar que $1/(\lambda - \alpha)$ es un valor propio de B, siendo $\mathbf{x}$ el correspondiente vector propio.

16. Suponga que μ es un valor propio de la B del ejercicio 15, y que $\mathbf{x}$ es el correspondiente vector propio, de manera que $(A - \alpha I)^{-1}\mathbf{x} = \mu\mathbf{x}$. Use esta ecuación para encontrar un valor propio de A en términos de μ y α. [*Nota:* $\mu \neq 0$ porque B es invertible.]

17. **[M]** Use el método de potencias inversas para estimar el valor propio medio de la A del ejemplo 3, con precisión de hasta cuatro posiciones decimales. Establezca $\mathbf{x}_0 = (1, 0, 0)$.

18. **[M]** Sea A como en el ejercicio 9. Utilice el método de potencias inversas con $\mathbf{x}_0 = (1, 0, 0)$ para estimar el valor propio de A cercano a $\alpha = -1.4$, con precisión de hasta cuatro posiciones decimales.

[M] En los ejercicios 19 y 20, encuentre (a) el valor propio más grande y (b) el valor propio más cercano a cero. En cada caso, establezca $\mathbf{x}_0 = (1, 0, 0, 0)$ y realice aproximaciones hasta que la sucesión aproximadora parezca tener una precisión de cuatro posiciones decimales. Incluya el vector propio aproximado.

19. $A = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}$

20. $A = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 12 & 13 & 11 \\ -2 & 3 & 0 & 2 \\ 4 & 5 & 7 & 2 \end{bmatrix}$

21. Una idea errónea común es que si A tiene un valor estrictamente dominante, entonces, para cualquier valor de k lo suficientemente grande, el vector $A^k\mathbf{x}$ es aproximadamente igual a un vector propio de A. Para las tres matrices siguientes, estudie qué le sucede a $A^k\mathbf{x}$ cuando $\mathbf{x} = (.5, .5)$, y trate de obtener conclusiones generales (para una matriz de 2×2).

a. $A = \begin{bmatrix} .8 & 0 \\ 0 & .2 \end{bmatrix}$

b. $A = \begin{bmatrix} 1 & 0 \\ 0 & .8 \end{bmatrix}$

c. $A = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$

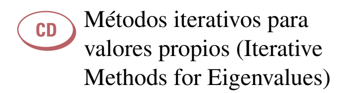
Solución al problema de práctica

Para las A y $\mathbf{x}$ dadas,

$$A\mathbf{x} = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1.00 \\ -4.30 \\ 8.10 \end{bmatrix} = \begin{bmatrix} 3.00 \\ -13.00 \\ 24.50 \end{bmatrix}$$

Si $A\mathbf{x}$ es casi un múltiplo de $\mathbf{x}$, entonces los cocientes de entradas correspondientes en los dos vectores deberían ser casi constantes. Así que calcule:

{ entrada en $A\mathbf{x}$ }	÷ { entrada en $\mathbf{x}$ }	= { cociente }
3.00	1.00	3.000
−13.00	−4.30	3.023
24.50	8.10	3.025



Métodos iterativos para valores propios (Iterative Methods for Eigenvalues)

Cada entrada de $A\mathbf{x}$ es cerca de 3 veces la entrada correspondiente de $\mathbf{x}$, así que $\mathbf{x}$ es cercano a un vector propio. Cualesquiera de los cocientes anteriores es una estimación del valor propio. (Hasta cinco posiciones decimales, el valor propio es 3.02409.)

Capítulo 5 Ejercicios suplementarios

En todos estos ejercicios suplementarios, A y B representan matrices cuadradas del tamaño apropiado.

1. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

a. Si A es invertible y 1 es un valor propio de A, entonces 1 también es valor propio de A^{-1}.

b. Si A es equivalente por filas a la matriz identidad I, entonces A es diagonalizable.

c. Si A contiene una columna o una fila de ceros, entonces 0 es un valor propio de A.

d. Cada valor propio de A también es un valor propio de A^2.

e. Cada vector propio de A también es un vector propio de A^2.

f. Cada vector propio de una matriz invertible A también es un vector propio de A^{-1}.

g. Los valores propios deben ser escalares diferentes de cero.

h. Los vectores propios deben ser vectores diferentes de cero.

i. Dos vectores propios, correspondientes al mismo valor propio, siempre son linealmente dependientes.

j. Las matrices semejantes siempre tienen exactamente los mismos valores propios.

k. Las matrices semejantes siempre tienen exactamente los mismos vectores propios.

l. La suma de dos vectores propios de una matriz A también es un vector propio de A.

m. Los valores propios de una matriz triangular superior A son exactamente las entradas diferentes de cero sobre la diagonal de A.

n. Las matrices A y A^T tienen los mismos valores propios, contando las multiplicidades.

o. Si una matriz A de 5×5 tiene menos de 5 valores propios distintos, entonces A no es diagonalizable.

p. Existe una matriz de 2×2 que no tiene vectores propios en $\mathbb{R}^2$.

q. Si A es diagonalizable, entonces las columnas de A son linealmente independientes.

r. Un vector diferente de cero no puede corresponder a dos diferentes valores propios de A.

s. Una matriz (cuadrada) A es invertible si, y sólo si, hay un sistema de coordenadas en el cual la transformación $\mathbf{x} \mapsto A\mathbf{x}$ esté representada por una matriz diagonal.

t. Si cada vector $\mathbf{e}_j$ en la base estándar para $\mathbb{R}^n$ es un vector propio de A, entonces A es una matriz diagonal.

u. Si A es semejante a una matriz diagonalizable B, entonces A también es diagonalizable.

v. Si A y B son matrices invertibles de $n \times n$, entonces AB es similar a BA.

w. Una matriz de $n \times n$, con n vectores propios linealmente independientes, es invertible.

x. Si A es una matriz diagonalizable de $n \times n$, entonces cada vector en $\mathbb{R}^n$ puede escribirse como una combinación lineal de vectores propios de A.

2. Muestre que si $\mathbf{x}$ es un vector propio del producto de matrices AB y $B\mathbf{x} \neq \mathbf{0}$, entonces $B\mathbf{x}$ es un vector propio de BA.

3. Suponga que $\mathbf{x}$ es un vector propio de A correspondiente a un valor propio λ.

a. Muestre que $\mathbf{x}$ es un vector propio de $5I - A$. ¿Cuál es el valor propio correspondiente?

b. Muestre que $\mathbf{x}$ es un vector propio de $5I - 3A + A^2$. ¿Cuál es el valor propio correspondiente?

4. Use inducción matemática para mostrar que si λ es un valor propio de una matriz A de $n \times n$, con $\mathbf{x}$ como el correspondiente vector propio, entonces, para cada entero positivo m, λ^m es un valor propio de A^m, siendo $\mathbf{x}$ un vector propio correspondiente.

5. Si $p(t) = c_0 + c_1t + c_2t^2 + \cdots + c_nt^n$, defina $p(A)$ como la matriz formada al reemplazar cada potencia de t en $p(t)$ por la potencia correspondiente de A (con $A^0 = I$). Es decir,

$$p(A) = c_0I + c_1\mathrm{A} + c_2\mathrm{A}^2 + \cdots + c_nA^n$$

Muestre que si λ es un valor propio de A, entonces un valor propio de $p(A)$ es $p(\lambda)$.

6. Suponga que $A = PDP^{-1}$, donde P es de 2×2 y $D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$.

a. Sea $B = 5I - 3A + A^2$. Muestre que B es diagonalizable encontrándole una factorización adecuada.

b. Dadas $p(t)$ y $p(A)$ como en el ejercicio 5, muestre que $p(A)$ es diagonalizable.

7. Suponga que A es diagonalizable y que $p(t)$ es el polinomio característico de A. Defina $p(A)$ como en el ejercicio 5, y muestre que $p(A)$ es la matriz cero. Este hecho, que es cierto para *cualquier* matriz cuadrada, se llama *teorema de Cayley-Hamilton.*

8. a. Sea A una matriz diagonalizable de $n \times n$. Muestre que si la multiplicidad de un valor propio λ es n, entonces $A = \lambda I$.

b. Use el inciso (a) para mostrar que la matriz $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ no es diagonalizable.

9. Muestre que $I - A$ es invertible cuando todos los valores propios de A son de magnitud menor que 1. [*Sugerencia:* ¿Qué sería cierto si $I - A$ no fuera invertible?]

10. Demuestre que si A es diagonalizable, con todos los valores propios de magnitud menor que 1, entonces A^k tiende a la matriz cero cuando $k \to \infty$. [*Sugerencia:* Considere $A^k\mathbf{x}$ donde $\mathbf{x}$ representa cualesquiera de las columnas de I.]

11. Sea $\mathbf{u}$ un vector propio de A correspondiente a un valor propio λ, y sea H la línea en $\mathbb{R}^n$ que pasa por $\mathbf{u}$ y el origen.

a. Explique por qué H es invariante bajo A en el sentido de que $A\mathbf{x}$ está en H siempre que $\mathbf{x}$ está en H.

b. Sea K un subespacio unidimensional de $\mathbb{R}^n$ que es invariante bajo A. Explique por qué K contiene un vector propio de A.

12. Sea $G = \begin{bmatrix} A & X \\ 0 & B \end{bmatrix}$. Use la fórmula (1) para el determinante de la sección 5.2 para explicar por qué det G = (det A)(det B). De ello, deduzca que el polinomio característico de G es el producto de los polinomios característicos de A y B.

Utilice el ejercicio 12 para encontrar los valores propios de las matrices:

13. $A = \begin{bmatrix} 3 & -2 & 8 \\ 0 & 5 & -2 \\ 0 & -4 & 3 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 5 & -6 & -7 \\ 2 & 4 & 5 & 2 \\ 0 & 0 & -7 & -4 \\ 0 & 0 & 3 & 1 \end{bmatrix}$

15. Sea J una matriz de $n \times n$ de sólo números 1, y considere $A = (a - b)I + bJ$; esto es,

$$A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Use los resultados del ejercicio 16 en los ejercicios suplementarios del capítulo 3 para mostrar que los valores de A son $a - b$ y $a + (n - 1)b$. ¿Cuáles son las multiplicidades de estos valores propios?

16. Aplique el resultado del ejercicio 15 para encontrar los valores propios de las matrices $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$ y

$$\begin{bmatrix} 7 & 3 & 3 & 3 & 3 \\ 3 & 7 & 3 & 3 & 3 \\ 3 & 3 & 7 & 3 & 3 \\ 3 & 3 & 3 & 7 & 3 \\ 3 & 3 & 3 & 3 & 7 \end{bmatrix}.$$

17. Sea $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Recuerde del ejercicio 25 en la sección 5.4 que tr A (la traza de A) es la suma de las entradas diagonales en A. Muestre que el polinomio característico de A es

$$\lambda^2 - (\text{tr } A)\lambda + \det A$$

Luego muestre que los valores propios de una matriz A de 2×2 son ambos reales si, y sólo si, $\det A \leq \left(\frac{\text{tr } A}{2}\right)^2$.

18. Sea $A = \begin{bmatrix} .4 & -.3 \\ .4 & 1.2 \end{bmatrix}$. Explique por qué A^k tiende a $\begin{bmatrix} -.5 & -.75 \\ 1.0 & 1.50 \end{bmatrix}$ conforme $k \to \infty$.

Los ejercicios 19 a 23 se refieren al polinomio

$$p(t) = a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + t^n$$

y a una matriz C_p de $n \times n$ llamada **matriz compañera** de p:

$$C_p = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}$$

19. Escriba la matriz compañera C_p de $p(t) = 6 - 5t + t^2$, y luego encuentre el polinomio característico de C_p.

20. Sea $p(t) = (t - 2)(t - 3)(t - 4) = -24 + 26t - 9t^2 + t^3$. Escriba la matriz compañera de $p(t)$ y use las técnicas del capítulo 3 para encontrar su polinomio característico.

21. Use inducción matemática y demuestre que, para $n \geq 2$,

$$\begin{aligned} \det(C_p - \lambda I) &= (-1)^n (a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n) \\ &= (-1)^n p(\lambda) \end{aligned}$$

[*Sugerencia:* Demuestre mediante el desarrollo por cofactores descendiendo por la primera columna que $\det(C_p - \lambda I)$ tiene la forma $(-\lambda)B + (-1)^n a_0$, donde B es cierto polinomio (por el supuesto de inducción).]

22. Sea $p(t) = a_0 + a_1 t + a_2 t^2 + t^3$, y sea λ un cero de p.

a. Escriba la matriz compañera de p.

b. Explique por qué $\lambda^3 = -a_0 - a_1\lambda - a_2\lambda^2$, y demuestre que $(1, \lambda, \lambda^2)$ es un vector propio de la matriz compañera de p.

23. Sea p el polinomio del ejercicio 22, y suponga que la ecuación $p(t) = 0$ tiene raíces distintas $\lambda_1, \lambda_2, \lambda_3$. Sea V la matriz de Vandermonde

$$V = \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{bmatrix}$$

(La transpuesta de V se consideró en el ejercicio suplementario 11 del capítulo 2.) Utilice el ejercicio 22 y un teorema de este capítulo para deducir que V es invertible (pero no calcule V^{-1}). Luego explique por que $V^{-1}C_pV$ es una matriz diagonal.

24. El comando `roots (p)` de MATLAB calcula las raíces de la ecuación polinomial $p(t) = 0$. Lea el manual de MATLAB y luego describa la idea en que se basa el algoritmo para el comando `roots`.

25. **[M]** Use un programa de matrices para diagonalizar

$$A = \begin{bmatrix} -3 & -2 & 0 \\ 14 & 7 & -1 \\ -6 & -3 & 1 \end{bmatrix}$$

si es posible. Use el comando de valores propios para crear la matriz diagonal D. Si el programa tiene un comando que produzca vectores propios, úselo para crear una matriz invertible P. Después calcule $AP - PD$ y PDP^{-1}. Analice sus resultados.

26. **[M]** Repita el ejercicio 25 para

$$A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}.$$

6

Ortogonalidad y mínimos cuadrados

EJEMPLO INTRODUCTORIO

Reajuste del Nivel de Referencia Norteamericano

Imagine comenzar un proyecto imponente que se estima tomará diez años concluir y requiere el esfuerzo de decenas de personas para estructurar y resolver un sistema de 1,800,000 por 900,000 ecuaciones lineales. Esto es exactamente lo que se hizo en 1974 durante el llamado Sondeo Geodésico Nacional, cuando se propuso actualizar el Nivel de Referencia Norteamericano (NAD, por sus siglas en inglés) —una red con 268,000 puntos de referencia cuidadosamente medidos y marcados que cubren todo el territorio de América del Norte situado al norte del Istmo de Panamá, junto con Groenlandia, Hawai, las Islas Vírgenes, Puerto Rico y otras islas del Caribe.

Las latitudes y longitudes registradas en el NAD deben determinarse con exactitud de unos pocos centímetros puesto que forman la base para trazar todos los planos, mapas, límites legales de la propiedad, fronteras estatales y regionales, y organizar proyectos de ingeniería civil como carreteras y líneas públicas de transmisión de electricidad. Desde el último ajuste —de los puntos de referencia geodésicos realizado en 1927, más de 200,000 nuevos puntos habían sido añadidos a un viejo conjunto de mediciones. Los errores se fueron acumulando gradualmente a través de los años, y en algunos lugares el terreno mismo se ha desplazado (hasta 5 centímetros por año). Hacia 1970 ya era urgente reacondicionar el sistema por completo y se hicieron planes para determinar un nuevo conjunto de coordenadas para los puntos de referencia.

Los datos de mediciones recopilados a lo largo de un periodo de 140 años debían convertirse a una forma legible por computadora, y los propios datos tenían que estandarizarse. (Por ejemplo, se usaron modelos matemáticos de los movimientos de la corteza terrestre para actualizar las mediciones efectuadas años atrás a lo largo de la falla de San Andrés en California.) Después de eso, había que comparar las mediciones para identificar errores surgidos de los datos originales o de los introducidos en la computadora. Los cálculos finales comprendían aproximadamente 1.8 millones de observaciones, ponderadas según su precisión relativa y cada una dando lugar a una ecuación.

El sistema de ecuaciones del NAD no tenía solución en el sentido común, pero sí una *solución por mínimos cuadrados*, la cual asignaba latitudes y longitudes a los puntos de referencia de tal modo que correspondieran de la mejor manera posible a los 1.8 millones de observaciones. Se encontró la solución de mínimos

cuadrados al resolver un sistema lineal relacionado de *ecuaciones normales*, el cual incluía 928,735 ecuaciones con 928,735 variables.

Como las ecuaciones normales eran demasiado grandes para las computadoras existentes, se descompusieron en sistemas más pequeños mediante una técnica llamada *bloqueo de Helmert*, la cual partía de manera recursiva la matriz de coeficientes en bloques cada vez más pequeños. Los bloques menores proporcionaban ecuaciones para bloques geográficamente contiguos de 500 a 2000 puntos de referencia del NAD. En la figura 1 se muestra cómo se subdividió Estados Unidos para conformar estos bloques de Helmert. Luego de algunos pasos intermedios se utilizaron las soluciones de los sistemas más pequeños para producir los valores finales de todas las 928,735 variables.[1]

En 1983 se completó la base de datos para el reajuste del NAD. Tres años después, tras un análisis extenso y más de 940 horas de procesamiento en computadora, se resolvió el mayor problema de mínimos cuadrados jamás intentado.

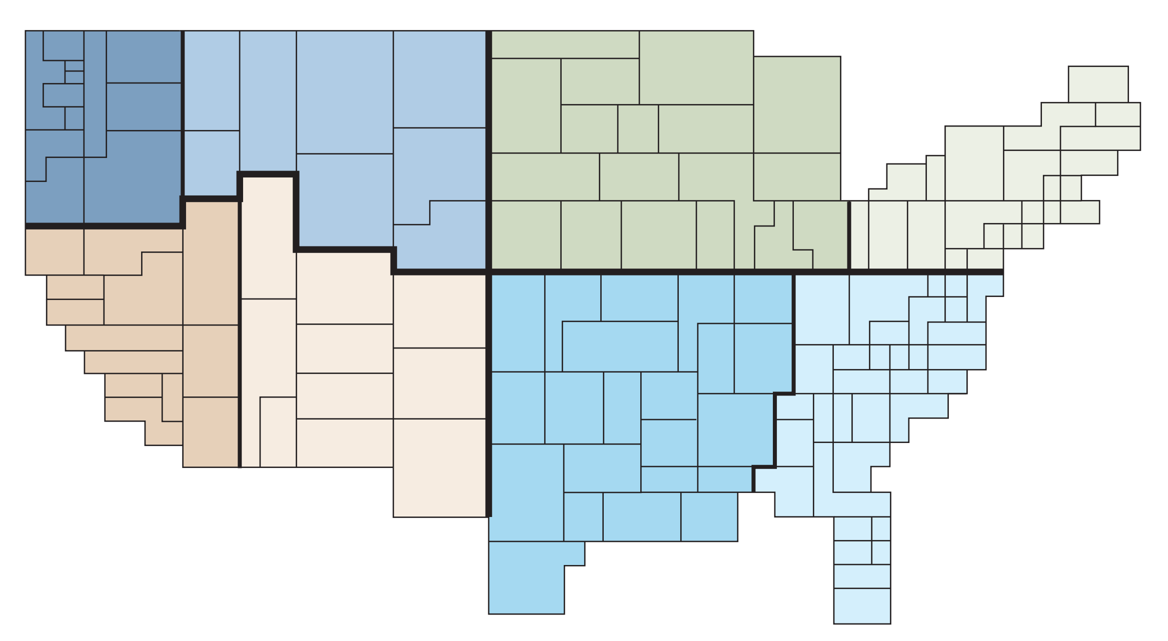


FIGURA 1 Fronteras de los bloques de Helmert contiguos para Estados Unidos.

Un sistema lineal $A\mathbf{x} = \mathbf{b}$ que surge de datos experimentales a menudo no tiene solución, como en el ejemplo introductorio. Con frecuencia, un sustituto aceptable de una solución es un vector $\hat{\mathbf{x}}$ que reduce la distancia entre $A\hat{\mathbf{x}}$ y $\mathbf{b}$ lo más posible. La definición de distancia, dada en la sección 6.1, involucra una suma de

[1]Un análisis matemático de la estrategia de bloques de Helmert, así como detalles acerca de todo el proyecto, aparecen en *North American Datum of 1983.* Charles R. Schwarz (ed.), National Geodetic Survey, National Oceanic and Atmospheric Administration (NOAA) Professional Paper NOS 2, 1989.

cuadrados, y la $\hat{\mathbf{x}}$ deseada se llama *solución de mínimos cuadrados* de $A\mathbf{x} = \mathbf{b}$. En las secciones 6.1, 6.2 y 6.3 se desarrollan los conceptos fundamentales de ortogonalidad y proyecciones ortogonales, los cuales se usan en la sección 6.5 para encontrar $\hat{\mathbf{x}}$.

En la sección 6.4 se proporciona otra oportunidad de observar el funcionamiento de las proyecciones ortogonales, al crear una factorización de matrices ampliamente usada en el álgebra lineal numérica. Las secciones restantes examinan algunos de los múltiples problemas de mínimos cuadrados que surgen en las aplicaciones, incluidos aquellos de espacios vectoriales más generales que $\mathbb{R}^n$. Sin embargo, en todos los casos los escalares son números reales.

6.1 PRODUCTO INTERIOR, LONGITUD Y ORTOGONALIDAD

Los conceptos geométricos de longitud, distancia y perpendicularidad, que son bien conocidos para $\mathbb{R}^2$ y $\mathbb{R}^3$, se definen aquí para $\mathbb{R}^n$. Estos conceptos proporcionan potentes herramientas geométricas para resolver muchos problemas aplicados, incluidos los problemas de mínimos cuadrados que ya se mencionaron. Las tres nociones se definen en términos del producto interior de dos vectores.

El producto interior

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$, entonces $\mathbf{u}$ y $\mathbf{v}$ se consideran como matrices de $n \times 1$. La transpuesta $\mathbf{u}^T$ es una matriz de $1 \times n$ y el producto matricial $\mathbf{u}^T\mathbf{v}$ es una matriz de 1×1, la cual se escribe como un solo número real (un escalar) sin corchetes. Al número $\boldsymbol{u}^T\mathbf{v}$ se le llama **producto interior** de $\mathbf{u}$ y $\mathbf{v}$, y se escribe a menudo como $\mathbf{u} \cdot \mathbf{v}$. Este producto interior, mencionado en los ejercicios de la sección 2.1, también se conoce como **producto punto**. Si

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

entonces el producto interior de $\mathbf{u}$ y $\mathbf{v}$ es

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

EJEMPLO 1 Calcule $\mathbf{u} \cdot \mathbf{v}$ y $\mathbf{v} \cdot \mathbf{u}$ cuando $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

Solución

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = [\,2 \quad -5 \quad -1\,]\begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

$$\mathbf{v}\cdot\mathbf{u} = \mathbf{v}^T\mathbf{u} = [\,3 \quad 2 \quad -3\,]\begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = (3)(2) + (2)(-5) + (-3)(-1) = -1$$

Por los cálculos del ejemplo 1, resulta claro por qué $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$. Esta conmutatividad del producto interior se aplica en lo general. Las siguientes propiedades del producto interior se deducen fácilmente de las propiedades de la operación transpuesta estudiada en la sección 2.1. (Vea los ejercicios 21 y 22 al final de esta sección.)

TEOREMA 1

Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores en $\mathbb{R}^n$, y sea c un escalar. Entonces,

a. $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$
b. $(\mathbf{u}+\mathbf{v})\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{w} + \mathbf{v}\cdot\mathbf{w}$
c. $(c\mathbf{u})\cdot\mathbf{v} = c(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot(c\mathbf{v})$
d. $\mathbf{u}\cdot\mathbf{u} \geq 0$, y $\mathbf{u}\cdot\mathbf{u} = 0$ si, y sólo si, $\mathbf{u} = \mathbf{0}$

Las propiedades (b) y (c) pueden combinarse varias veces para producir la siguiente regla útil:

$$(c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{w} = c_1(\mathbf{u}_1\cdot\mathbf{w}) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{w})$$

La longitud de un vector

Si $\mathbf{v}$ está en $\mathbb{R}^n$, con entradas $v_1, \ldots, v_n$, entonces la raíz cuadrada de $\mathbf{v}\cdot\mathbf{v}$ está definida porque $\mathbf{v}\cdot\mathbf{v}$ no es negativo.

DEFINICIÓN

La **longitud** (o **norma)** de $\mathbf{v}$ es el escalar no negativo $\|\mathbf{v}\|$ definido mediante

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{y} \quad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$$

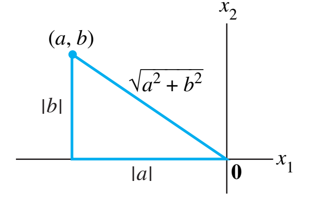


FIGURA 1 Interpretación de $\|\mathbf{v}\|$ como longitud.

Suponga que $\mathbf{v}$ está en $\mathbb{R}^2$, por ejemplo, $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. Si se identifica a $\mathbf{v}$ con un punto geométrico en el plano, como siempre, entonces $\|\mathbf{v}\|$ coincide con la noción estándar de la longitud del segmento de línea que va desde el origen hasta $\mathbf{v}$. Esto se deriva del teorema de Pitágoras aplicado a un triángulo como el de la figura 1.

Un cálculo similar con la diagonal de una caja rectangular muestra que la definición de la longitud de un vector $\mathbf{v}$ en $\mathbb{R}^3$ coincide con la noción usual de longitud.

Para cualquier escalar c, la longitud de $c\mathbf{v}$ es $|c|$ veces la longitud de $\mathbf{v}$. Esto es,

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$$

(Para ver esto, calcule $\|c\mathbf{v}\|^2 = (c\mathbf{v})\cdot(c\mathbf{v}) = c^2\mathbf{v}\cdot\mathbf{v} = c^2\|\mathbf{v}\|^2$ y obtenga las raíces cuadradas.)

Un vector cuya longitud es 1 se llama **vector unitario**. Si se *divide* un vector $\mathbf{v}$ diferente de cero entre su longitud —esto es, se multiplica por $1/\|\mathbf{v}\|$— se obtiene un vector unitario $\mathbf{u}$ porque la longitud de $\mathbf{u}$ es $(1/\|\mathbf{v}\|)\|\mathbf{v}\|$. El proceso de crear a $\mathbf{u}$ a partir de $\mathbf{v}$ en ocasiones se denomina **normalización de v**, y se dice que $\mathbf{u}$ está *en la misma dirección que* $\mathbf{v}$.

Varios de los ejemplos siguientes utilizan una notación con la que se ahorra espacio con vectores (columna).

EJEMPLO 2 Sea $\mathbf{v} = (1, -2, 2, 0)$. Encuentre un vector unitario $\mathbf{u}$ en la misma dirección que $\mathbf{v}$.

Solución Primero, determine la longitud de $\mathbf{v}$:

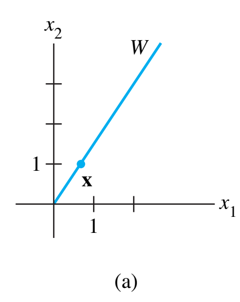
$$\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v} = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9$$
$$\|\mathbf{v}\| = \sqrt{9} = 3$$

Después, multiplique $\mathbf{v}$ por $1/\|\mathbf{v}\|$ para obtener

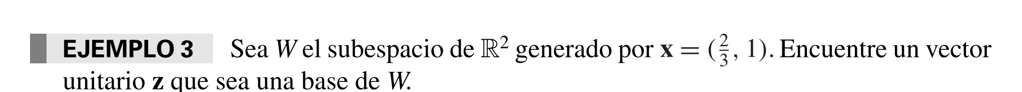
$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{3}\mathbf{v} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$$

Para comprobar que $\|\mathbf{u}\| = 1$, basta con demostrar que $\|\mathbf{u}\|^2 = 1$.

$$\|\mathbf{u}\|^2 = \mathbf{u}\cdot\mathbf{u} = \left(\tfrac{1}{3}\right)^2 + \left(-\tfrac{2}{3}\right)^2 + \left(\tfrac{2}{3}\right)^2 + (0)^2$$
$$= \tfrac{1}{9} + \tfrac{4}{9} + \tfrac{4}{9} + 0 = 1$$



(a)



EJEMPLO 3 Sea W el subespacio de $\mathbb{R}^2$ generado por $\mathbf{x} = (\frac{2}{3}, 1)$. Encuentre un vector unitario $\mathbf{z}$ que sea una base de W.

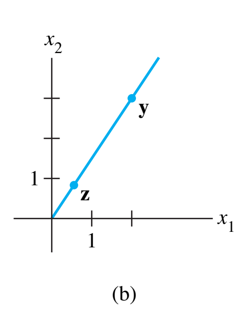
Solución W consta de todos los múltiplos de $\mathbf{x}$, como en la figura 2(a). Cualquier vector en W que sea diferente de cero es una base de W. Para simplificar el cálculo, "escale" $\mathbf{x}$ para eliminar las fracciones. Esto es, se multiplica $\mathbf{x}$ por 3 para obtener

$$\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Ahora calcule $\|\mathbf{y}\|^2 = 2^2 + 3^2 = 13$, $\|\mathbf{y}\| = \sqrt{13}$, y normalice $\mathbf{y}$ para obtener

$$\mathbf{z} = \frac{1}{\sqrt{13}}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}$$

Vea la figura 2(b). Otro vector unitario es $(-2/\sqrt{13}, -3/\sqrt{13})$.



(b)

FIGURA 2
Normalización de un vector para producir un vector unitario.

Distancia en $\mathbb{R}^n$

Ahora se tiene la capacidad de describir qué tan cercano es un vector a otro. Recuerde que si a y b son números reales, la distancia sobre la recta numérica entre a y b es el número $|a - b|$. En la figura 3, se muestran dos ejemplos. Esta definición de distancia en $\mathbb{R}$ tiene un análogo directo en $\mathbb{R}^n$.

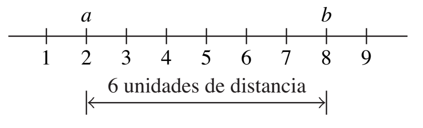


FIGURA 3 Distancias en $\mathbb{R}$.

DEFINICIÓN

Para $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, la **distancia entre u y v**, escrita como dist($\mathbf{u}$, $\mathbf{v}$), es la longitud del vector $\mathbf{u} - \mathbf{v}$. Esto es,

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

En $\mathbb{R}^2$ y $\mathbb{R}^3$, esta definición de distancia coincide con las fórmulas usuales para la distancia euclidiana entre dos puntos, tal como indican los dos ejemplos siguientes.

EJEMPLO 4 Encuentre la distancia entre los vectores $\mathbf{u} = (7, 1)$ y $\mathbf{v} = (3, 2)$.

Solución Calcule

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

Los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ se muestran en la figura 4. Cuando se suma el vector $\mathbf{u} - \mathbf{v}$ a $\mathbf{v}$, el resultado es $\mathbf{u}$. Observe que el paralelogramo de la figura 4 muestra que la distancia de $\mathbf{v}$ a $\mathbf{u}$ es la misma que de $\mathbf{u} - \mathbf{v}$ a $\mathbf{0}$.

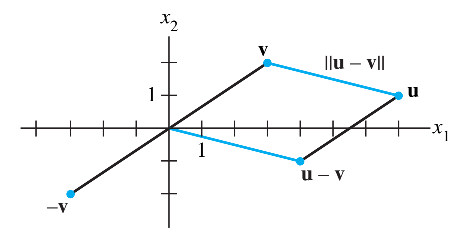


FIGURA 4 La distancia entre $\mathbf{u}$ y $\mathbf{v}$ es la longitud de $\mathbf{u} - \mathbf{v}$.

EJEMPLO 5 Si $\mathbf{u} = (u_1, u_2, u_3)$ y $\mathbf{v} = (v_1, v_2, v_3)$, entonces

$$\begin{aligned}\text{dist}(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\mathbf{u} - \mathbf{v})\cdot(\mathbf{u} - \mathbf{v})} \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}\end{aligned}$$

Vectores ortogonales

El resto de este capítulo se subordina al hecho de que el concepto de líneas perpendiculares en la geometría euclidiana ordinaria tiene un análogo en $\mathbb{R}^n$.

Considere $\mathbb{R}^2$ o $\mathbb{R}^3$ y dos líneas que pasen por el origen determinadas mediante los vectores **u** y **v**. Las dos líneas que se muestran en la figura 5 son geométricamente perpendiculares si, y sólo si, la distancia desde **u** hasta **v** es igual a la distancia desde **u** hasta $-\mathbf{v}$. Esto es análogo a pedir que los cuadrados de las distancias sean iguales. Ahora

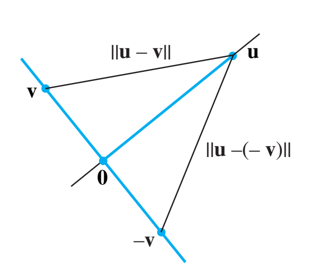


FIGURA 5

$$\begin{aligned}[\,\text{dist}(\mathbf{u}, -\mathbf{v})\,]^2 &= \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \\ &= (\mathbf{u} + \mathbf{v})\cdot(\mathbf{u} + \mathbf{v}) \\ &= \mathbf{u}\cdot(\mathbf{u} + \mathbf{v}) + \mathbf{v}\cdot(\mathbf{u} + \mathbf{v}) && \text{Teorema 1(b)} \\ &= \mathbf{u}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u} + \mathbf{v}\cdot\mathbf{v} && \text{Teorema 1(a), (b)} \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u}\cdot\mathbf{v} && \text{Teorema 1(a)} \qquad (1)\end{aligned}$$

Los mismos cálculos con **v** y $-\mathbf{v}$ intercambiados muestran que

$$\begin{aligned}[\text{dist}(\mathbf{u}, \mathbf{v})]^2 &= \|\mathbf{u}\|^2 + \|-\mathbf{v}\|^2 + 2\mathbf{u}\cdot(-\mathbf{v}) \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}\cdot\mathbf{v}\end{aligned}$$

Las dos distancias elevadas al cuadrado son iguales si, y sólo si, $2\mathbf{u}\cdot\mathbf{v} = -2\mathbf{u}\cdot\mathbf{v}$, lo cual sucede si, y sólo si, $\mathbf{u}\cdot\mathbf{v} = 0$.

Estos cálculos muestran que cuando los vectores **u** y **v** se identifican con puntos geométricos, las líneas correspondientes que pasan por los puntos y el origen son perpendiculares si, y sólo si, $\mathbf{u}\cdot\mathbf{v} = 0$. La siguiente definición generaliza a $\mathbb{R}^n$ esta noción de perpendicularidad (u *ortogonalidad*, como se le llama comúnmente en álgebra lineal).

DEFINICIÓN Dos vectores **u** y **v** en $\mathbb{R}^n$ son **ortogonales** (entre sí) si $\mathbf{u}\cdot\mathbf{v} = 0$.

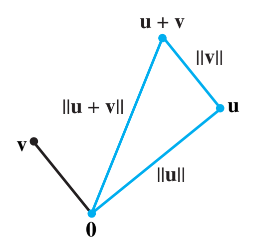


FIGURA 6

Observe que el vector cero es ortogonal a todo vector en $\mathbb{R}^n$ porque $\mathbf{0}^T\mathbf{v} = 0$ para toda **v**.

En el teorema siguiente se proporciona un dato clave acerca de los vectores ortogonales. La demostración se deriva inmediatamente a partir de los cálculos incluidos en (1) antes de la definición de ortogonalidad. El triángulo rectángulo mostrado en la figura 6 proporciona una visualización de las longitudes que aparecen en el teorema.

TEOREMA 2 **El teorema de Pitágoras**

Dos vectores $\mathbf{u}$ y $\mathbf{v}$ son ortogonales si, y sólo si, $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

Complementos ortogonales

Con la intención de que el lector pueda practicar el uso de los productos interiores, se introduce aquí un concepto que será útil en la sección 6.3 y en el resto del capítulo. Si un vector $\mathbf{z}$ es ortogonal a todo vector en un subespacio W de $\mathbb{R}^n$, se afirma entonces que $\mathbf{z}$ es **ortogonal a** W. El conjunto de todos los vectores $\mathbf{z}$ que son ortogonales a W se llama **complemento ortogonal** de W y se denota mediante $W^\perp$ (se lee "W perpendicular o simplemente "W perp").

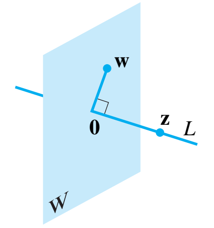


FIGURA 7
Un plano y la línea que pasa por **0** como complementos ortogonales.

EJEMPLO 6 Sea W un plano que pasa por el origen en $\mathbb{R}^3$, y sea L la línea que pasa por el origen y es perpendicular a W. Si $\mathbf{z}$ y $\mathbf{w}$ son diferentes de cero, $\mathbf{z}$ está sobre L, y $\mathbf{w}$ está en W, entonces el segmento de línea de $\mathbf{0}$ a $\mathbf{z}$ es perpendicular al segmento de línea de $\mathbf{0}$ a $\mathbf{w}$; esto es, $\mathbf{z}\cdot\mathbf{w} = 0$. Vea la figura 7. Así que cada vector sobre L es ortogonal a cada $\mathbf{w}$ en W. De hecho, L consiste en *todos* los vectores que son ortogonales a los $\mathbf{w}$ en W, y W consiste en todos los vectores ortogonales a los vectores $\mathbf{z}$ en L. Esto es,

$$L = W^\perp \quad \text{y} \quad W = L^\perp$$

Los dos hechos siguientes acerca de $W^\perp$, siendo W un subespacio de $\mathbb{R}^n$, se necesitarán más adelante en el capítulo. Las demostraciones se sugieren en los ejercicios 29 y 30. Los ejercicios del 27 al 31 ofrecen una práctica excelente para el uso de las propiedades del producto interior.

1. Un vector $\mathbf{x}$ está en $W^\perp$ si, y sólo si, $\mathbf{x}$ es ortogonal a todo vector de un conjunto que genere a W.
2. $W^\perp$ es un subespacio de $\mathbb{R}^n$.

El siguiente teorema y el ejercicio 31 prueban las afirmaciones contenidas en la sección 4.6 acerca de los subespacios que se muestran en la figura 8. (Vea también el ejercicio 28 de la sección 4.6.)

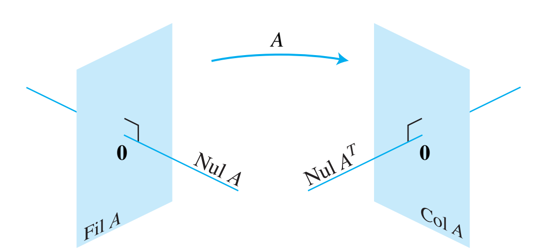


FIGURA 8 Los subespacios fundamentales determinados por una matriz A de $m \times n$.

TEOREMA 3

Sea A una matriz de $m \times n$. El complemento ortogonal del espacio fila de A es el espacio nulo de A, y el complemento ortogonal del espacio columna de A es el espacio nulo de A^T:

$$(\text{Fil } A)^{\perp} = \text{Nul } A \quad \text{y} \quad (\text{Col } A)^{\perp} = \text{Nul } A^T$$

DEMOSTRACIÓN La regla fila-columna para calcular $A\mathbf{x}$ muestra que si $\mathbf{x}$ está en Nul A, entonces $\mathbf{x}$ es ortogonal a cada fila de A (si las filas se tratan como vectores en $\mathbb{R}^n$). Dado que las filas de A generan el espacio fila, $\mathbf{x}$ es ortogonal a Fil A. De manera recíproca, si $\mathbf{x}$ es ortogonal a Fil A, entonces, evidentemente, $\mathbf{x}$ es ortogonal a cada fila de A y, por lo tanto, $A\mathbf{x} = \mathbf{0}$. Esto prueba la primera afirmación del teorema. Puesto que este enunciado es cierto para cualquier matriz, es verdadero también para A^T. Es decir, el complemento ortogonal del espacio fila de A^T es el espacio nulo de A^T. Lo cual demuestra el segundo enunciado, porque Fil A^T = Col A. ■

Ángulos en $\mathbb{R}^2$ y $\mathbb{R}^3$ (opcional)

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores diferentes de cero en $\mathbb{R}^2$ o en $\mathbb{R}^3$, entonces existe una conexión interesante entre sus productos interiores y el ángulo ϑ entre los dos segmentos de línea que van desde el origen hasta los puntos identificados con $\mathbf{u}$ y $\mathbf{v}$. La fórmula es

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos \vartheta \tag{2}$$

Para verificar esta fórmula para vectores en $\mathbb{R}^2$, considere el triángulo mostrado en la figura 9, cuyos lados tienen longitudes $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, y $\|\mathbf{u} - \mathbf{v}\|$. Por la ley de los cosenos,

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \, \|\mathbf{v}\| \cos \vartheta$$

lo cual puede reordenarse para obtener

$$\begin{aligned} \|\mathbf{u}\| \, \|\mathbf{v}\| \cos \vartheta &= \frac{1}{2}\left[\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2\right] \\ &= \frac{1}{2}\left[u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2\right] \\ &= u_1 v_1 + u_2 v_2 \\ &= \mathbf{u} \cdot \mathbf{v} \end{aligned}$$

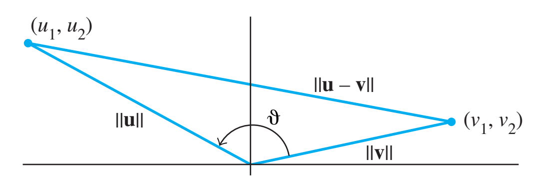


FIGURA 9 El ángulo entre dos vectores.

La verificación para $\mathbb{R}^3$ es semejante. Cuando $n > 3$, la fórmula (2) se puede usar para *definir* el ángulo entre dos vectores de $\mathbb{R}^n$. En estadística, por ejemplo, el valor de $\cos\vartheta$ definido mediante (2) para los vectores **u** y **v** es lo que los estadísticos llaman un *coeficiente de correlación*.

PROBLEMAS DE PRÁCTICA

Sean $\mathbf{a} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix}$, y $\mathbf{d} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix}$.

1. Calcule $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}$ y $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a}$.

2. Encuentre un vector unitario **u** en la dirección de **c**.

3. Muestre que **d** es ortogonal a **c**.

4. Use los resultados de los problemas de práctica 2 y 3 para explicar por qué **d** tiene que ser ortogonal al vector unitario **u**.

6.1 EJERCICIOS

En los ejercicios 1 a 8, determine las cantidades usando los vectores

$$\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix}$$

1. $\mathbf{u}\cdot\mathbf{u}$, $\mathbf{v}\cdot\mathbf{u}$, y $\dfrac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$

2. $\mathbf{w}\cdot\mathbf{w}$, $\mathbf{x}\cdot\mathbf{w}$ y $\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}$

3. $\dfrac{1}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}$

4. $\dfrac{1}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$

5. $\left(\dfrac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$

6. $\left(\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{x}\cdot\mathbf{x}}\right)\mathbf{x}$

7. $\|\mathbf{w}\|$

8. $\|\mathbf{x}\|$

En los ejercicios 9 a 12, encuentre un vector unitario en la dirección del vector dado.

9. $\begin{bmatrix} -30 \\ 40 \end{bmatrix}$

10. $\begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix}$

11. $\begin{bmatrix} 7/4 \\ 1/2 \\ 1 \end{bmatrix}$

12. $\begin{bmatrix} 8/3 \\ 2 \end{bmatrix}$

13. Encuentre la distancia entre $\mathbf{x} = \begin{bmatrix} 10 \\ -3 \end{bmatrix}$ y $\mathbf{y} = \begin{bmatrix} -1 \\ -5 \end{bmatrix}$.

14. Encuentre la distancia entre $\mathbf{u} = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}$ y $\mathbf{z} = \begin{bmatrix} -4 \\ -1 \\ 8 \end{bmatrix}$.

Determine cuáles pares de vectores en los ejercicios 15 al 18 son ortogonales

15. $\mathbf{a} = \begin{bmatrix} 8 \\ -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$

16. $\mathbf{u} = \begin{bmatrix} 12 \\ 3 \\ -5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix}$

17. $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ -5 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -4 \\ 1 \\ -2 \\ 6 \end{bmatrix}$

18. $\mathbf{y} = \begin{bmatrix} -3 \\ 7 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} 1 \\ -8 \\ 15 \\ -7 \end{bmatrix}$

En los ejercicios 19 y 20, todos los vectores están en $\mathbb{R}^n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

19. a. $\mathbf{v}\cdot\mathbf{v} = \|\mathbf{v}\|^2$.

b. Para cualquier escalar c, $\mathbf{u}\cdot(c\mathbf{v}) = c(\mathbf{u}\cdot\mathbf{v})$.

c. Si la distancia de **u** a **v** es igual a la distancia de **u** a $-\mathbf{v}$, entonces **u** y **v** son ortogonales.

d. Para una matriz cuadrada A, los vectores de Col A son ortogonales a los vectores de Nul A.

e. Si los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_p$ generan un subespacio W y si **x** es ortogonal a cada $\mathbf{v}_j$ para $j = 1, \ldots, p$, entonces **x** está en $W^\perp$.

20. a. $\mathbf{u}\cdot\mathbf{v} - \mathbf{v}\cdot\mathbf{u} = 0$.

b. Para cualquier escalar c, $\|c\mathbf{v}\| = c\|\mathbf{v}\|$.

c. Si **x** es ortogonal a todo vector en un subespacio W, entonces **x** está en $W^\perp$.

d. Si $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u} + \mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

e. Para una matriz A de $m \times n$, los vectores en el espacio nulo de A son ortogonales a los vectores en el espacio fila de A.

21. Use la definición de transpuesta del producto interior para verificar los incisos (b) y (c) del teorema 1. Mencione las referencias apropiadas del capítulo 2.

22. Sea $\mathbf{u} = (u_1, u_2, u_3)$. Explique por qué $\mathbf{u}\cdot\mathbf{u} \geq 0$. ¿Cuándo es $\mathbf{u}\cdot\mathbf{u} = 0$?

23. Sean $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -7 \\ -4 \\ 6 \end{bmatrix}$. Calcule y compare $\mathbf{u}\cdot\mathbf{v}$, $\|\mathbf{u}\|^2$, $\|\mathbf{v}\|^2$, y $\|\mathbf{u} + \mathbf{v}\|^2$. No utilice el teorema de Pitágoras.

24. Verifique la *ley del paralelogramo* para los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$:

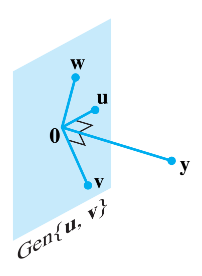
$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

25. Sea $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. Describa el conjunto H de vectores $\begin{bmatrix} x \\ y \end{bmatrix}$ que son ortogonales a $\mathbf{v}$. [*Sugerencia:* Considere $\mathbf{v} = \mathbf{0}$ y $\mathbf{v} \neq \mathbf{0}$.]

26. Sea $\mathbf{u} = \begin{bmatrix} 5 \\ -6 \\ 7 \end{bmatrix}$, y sea W el conjunto de todos los $\mathbf{x}$ en $\mathbb{R}^3$ tales que $\mathbf{u}\cdot\mathbf{x} = 0$. ¿Qué teorema del capítulo 4 puede usarse para demostrar que W es un subespacio de $\mathbb{R}^3$? Describa W en lenguaje geométrico.

27. Suponga que un vector $\mathbf{y}$ es ortogonal a los vectores $\mathbf{u}$ y $\mathbf{v}$. Muestre que $\mathbf{y}$ es ortogonal al vector $\mathbf{u} + \mathbf{v}$.

28. Suponga que $\mathbf{y}$ es ortogonal a $\mathbf{u}$ y $\mathbf{v}$. Muestre que $\mathbf{y}$ es ortogonal a todo $\mathbf{w}$ en Gen$\{\mathbf{u}, \mathbf{v}\}$. [*Sugerencia:* Un $\mathbf{w}$ arbitrario en Gen$\{\mathbf{u}, \mathbf{v}\}$ tiene la forma $\mathbf{w} = c_1\mathbf{u} + c_2\mathbf{v}$. Muestre que $\mathbf{y}$ es ortogonal a tal vector $\mathbf{w}$.]



29. Sea $W = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$. Muestre que si $\mathbf{x}$ es ortogonal a todo $\mathbf{v}_j$, para $1 \leq j \leq p$, entonces $\mathbf{x}$ es ortogonal a todo vector de W.

30. Sea W un subespacio de $\mathbb{R}^n$, y sea $W^\perp$ el conjunto de todos los vectores ortogonales a W. Muestre que $W^\perp$ es un subespacio de $\mathbb{R}^n$ usando los siguientes pasos:

a. Tome $\mathbf{z}$ en $W^\perp$, y sea $\mathbf{u}$ tal que represente a cualquier elemento de W. Entonces $\mathbf{z}\cdot\mathbf{u} = 0$. Tome cualquier escalar c y muestre que $c\mathbf{z}$ es ortogonal a $\mathbf{u}$. (Puesto que $\mathbf{u}$ era un elemento arbitrario de W, esto mostrará que $c\mathbf{z}$ está en $W^\perp$.)

b. Tome $\mathbf{z}_1$ y $\mathbf{z}_2$ de $W^\perp$, y sea $\mathbf{u}$ cualquier elemento de W. Muestre que $\mathbf{z}_1 + \mathbf{z}_2$ es ortogonal a $\mathbf{u}$. ¿Qué puede concluirse acerca de $\mathbf{z}_1 + \mathbf{z}_2$? ¿Por qué?

c. Termine la demostración de que $W^\perp$ es un subespacio de $\mathbb{R}^n$.

31. Muestre que si $\mathbf{x}$ está en W y en $W^\perp$, entonces $\mathbf{x} = \mathbf{0}$.

32. **[M]** Construya un par $\mathbf{u}$, $\mathbf{v}$ de vectores al azar en $\mathbb{R}^4$, y sea

$$A = \begin{bmatrix} .5 & .5 & .5 & .5 \\ .5 & .5 & -.5 & -.5 \\ .5 & -.5 & .5 & -.5 \\ .5 & -.5 & -.5 & .5 \end{bmatrix}$$

a. Denote las columnas de A mediante $\mathbf{a}_1, \ldots, \mathbf{a}_4$. Determine la longitud de cada columna y calcule $\mathbf{a}_1\cdot\mathbf{a}_2$, $\mathbf{a}_1\cdot\mathbf{a}_3$, $\mathbf{a}_1\cdot\mathbf{a}_4$, $\mathbf{a}_2\cdot\mathbf{a}_3$, $\mathbf{a}_2\cdot\mathbf{a}_4$ y $\mathbf{a}_3\cdot\mathbf{a}_4$.

b. Calcule y compare las longitudes de $\mathbf{u}$, $A\mathbf{u}$, $\mathbf{v}$, y $A\mathbf{v}$.

c. Utilice la ecuación (2) de esta sección para calcular el coseno del ángulo entre $\mathbf{u}$ y $\mathbf{v}$. Compárelo con el coseno del ángulo entre $A\mathbf{u}$ y $A\mathbf{v}$.

d. Repita los incisos (b) y (c) para otros dos pares de vectores al azar. ¿Qué conjeturas pueden formularse acerca del efecto de A sobre los vectores?

33. **[M]** Genere vectores al azar $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{v}$ en $\mathbb{R}^4$ con entradas enteras (y $\mathbf{v} \neq \mathbf{0}$), y determine las cantidades

$$\left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \left(\frac{\mathbf{y}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \frac{(\mathbf{x}+\mathbf{y})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v},\ \frac{(10\mathbf{x})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}$$

Repita los cálculos con nuevos vectores al azar $\mathbf{x}$ y $\mathbf{y}$. ¿Qué conjetura puede formularse acerca de la función $\mathbf{x} \mapsto T(\mathbf{x}) = \left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$ (para $\mathbf{v} \neq \mathbf{0}$)? Verifique su conjetura de manera algebraica.

34. **[M]** Sea $A = \begin{bmatrix} -6 & 3 & -27 & -33 & -13 \\ 6 & -5 & 25 & 28 & 14 \\ 8 & -6 & 34 & 38 & 18 \\ 12 & -10 & 50 & 41 & 23 \\ 14 & -21 & 49 & 29 & 33 \end{bmatrix}$. Construya una matriz N cuyas columnas formen una base para Nul A, y una matriz R cuyas *filas* formen una base para Fil A. (Consulte la sección 4.6 para ver mayores detalles.) Realice un cálculo matricial con N y R que ilustre un hecho del teorema 3.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. $\mathbf{a}\cdot\mathbf{b} = 7$, $\mathbf{a}\cdot\mathbf{a} = 5$. Por lo tanto, $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}} = \dfrac{7}{5}$, y $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a} = \dfrac{7}{5}\mathbf{a} = \begin{bmatrix} -14/5 \\ 7/5 \end{bmatrix}$.

2. Escale **c** al multiplicarlo por 3 para obtener $\mathbf{y} = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$. Calcule $\|\mathbf{y}\|^2 = 29$ y $\|\mathbf{y}\| = \sqrt{29}$.

   El vector unitario en la dirección tanto de **c** como de **y** es $\mathbf{u} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y} = \begin{bmatrix} 4/\sqrt{29} \\ -3/\sqrt{29} \\ 2/\sqrt{29} \end{bmatrix}$.

3. **d** es ortogonal a **c** porque

$$\mathbf{d}\cdot\mathbf{c} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix} = \frac{20}{3} - 6 - \frac{2}{3} = 0$$

4. **d** es ortogonal a **u** porque **u** tiene la forma $k\mathbf{c}$ para alguna k, y

$$\mathbf{d}\cdot\mathbf{u} = \mathbf{d}\cdot(k\mathbf{c}) = k(\mathbf{d}\cdot\mathbf{c}) = k(0) = 0$$

6.2 CONJUNTOS ORTOGONALES

Se dice que un conjunto de vectores $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ en $\mathbb{R}^n$ es un **conjunto ortogonal** si cada par de vectores distintos en el conjunto es ortogonal, esto es, si $\mathbf{u}_i\cdot\mathbf{u}_j = 0$ siempre que $i \neq j$.

EJEMPLO 1 Muestre que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal, donde

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

Solución Considere los tres pares posibles de vectores, es decir, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, y $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\begin{aligned} \mathbf{u}_1\cdot\mathbf{u}_2 &= 3(-1) + 1(2) + 1(1) = 0 \\ \mathbf{u}_1\cdot\mathbf{u}_3 &= 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0 \\ \mathbf{u}_2\cdot\mathbf{u}_3 &= -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0 \end{aligned}$$

Cada par de vectores distintos es ortogonal, así que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal. Vea la figura 1; los tres segmentos de línea que se muestran son mutuamente perpendiculares.

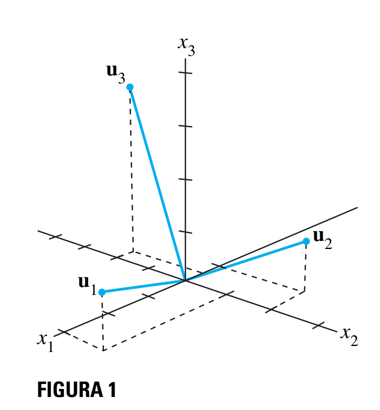


FIGURA 1

TEOREMA 4

Si $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es un conjunto ortogonal de vectores diferentes de cero en $\mathbb{R}^n$, entonces S es linealmente independiente y, por lo tanto, es una base del subespacio generado por S.

DEMOSTRACIÓN Si $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ para algunos escalares $c_1, \ldots, c_p$, entonces

$$\begin{aligned} 0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) \end{aligned}$$

porque $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_2, \ldots, \mathbf{u}_p$. Como $\mathbf{u}_1$ es diferente de cero, $\mathbf{u}_1\cdot\mathbf{u}_1$ no es cero y, por lo tanto, $c_1 = 0$. De manera similar, $c_2, \ldots, c_p$ deben ser cero. Así que S es linealmente independiente. ■

DEFINICIÓN Una **base ortogonal** para un subespacio W de $\mathbb{R}^n$ es una base para W que también es un conjunto ortogonal.

El teorema siguiente sugiere por qué una base ortogonal es mucho mejor que otras bases: Los pesos de una combinación lineal pueden calcularse fácilmente.

TEOREMA 5 Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ una base ortogonal para un subespacio W de $\mathbb{R}^n$. Para cada $\mathbf{y}$ en W, los pesos en la combinación lineal

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

están dados por

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \qquad (j = 1, \ldots, p)$$

DEMOSTRACIÓN Igual que en la demostración anterior, la ortogonalidad de $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ muestra que

$$\mathbf{y}\cdot\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1)$$

Como $\mathbf{u}_1\cdot\mathbf{u}_1$ no es cero, la ecuación anterior puede resolverse para c_1. Para encontrar c_j para $j = 2, \ldots, p$, se calcula $\mathbf{y}\cdot\mathbf{u}_j$ y se despeja c_j. ■

EJEMPLO 2 El conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ del ejemplo 1 es una base ortogonal para $\mathbb{R}^3$.

Exprese el vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ como una combinación lineal de los vectores en S.

Solución Calcule

$$\begin{aligned} \mathbf{y}\cdot\mathbf{u}_1 &= 11, & \mathbf{y}\cdot\mathbf{u}_2 &= -12, & \mathbf{y}\cdot\mathbf{u}_3 &= -33 \\ \mathbf{u}_1\cdot\mathbf{u}_1 &= 11, & \mathbf{u}_2\cdot\mathbf{u}_2 &= 6, & \mathbf{u}_3\cdot\mathbf{u}_3 &= 33/2 \end{aligned}$$

De acuerdo con el teorema 5,

$$\begin{aligned}\mathbf{y} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3 \\ &= \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-33}{33/2}\mathbf{u}_3 \\ &= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3\end{aligned}$$

Observe lo fácil que es calcular los pesos necesarios para construir **y** a partir de una base ortogonal. Si la base no fuera ortogonal, habría que resolver un sistema de ecuaciones lineales para poder encontrar los pesos, como en el capítulo 1.

Enseguida se verá una estructura que va a constituirse en paso clave para muchos de los cálculos que involucran ortogonalidad, y conducirá a una interpretación geométrica del teorema 5.

Una proyección ortogonal

Dado un vector **u** diferente de cero en $\mathbb{R}^n$, considere el problema de descomponer un vector **y** de $\mathbb{R}^n$ en la suma de dos vectores, uno un múltiplo de **u** y el otro ortogonal a **u**. Se desea escribir

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

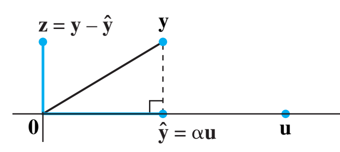


FIGURA 2
Cómo encontrar un α para hacer que $\mathbf{y} - \hat{\mathbf{y}}$ sea ortogonal a **u**.

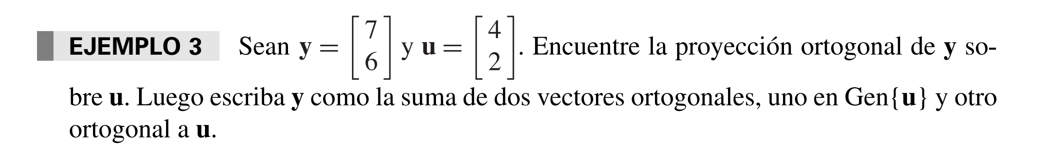
donde $\hat{\mathbf{y}} = \alpha\mathbf{u}$ para algún escalar α y **z** es algún vector ortogonal a **u**. Vea la figura 2. Dado cualquier escalar α, sea $\mathbf{z} = \mathbf{y} - \alpha\mathbf{u}$, de manera que (1) se cumple. Entonces $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a **u** si, y sólo si,

$$0 = (\mathbf{y} - \alpha\mathbf{u})\cdot\mathbf{u} = \mathbf{y}\cdot\mathbf{u} - (\alpha\mathbf{u})\cdot\mathbf{u} = \mathbf{y}\cdot\mathbf{u} - \alpha(\mathbf{u}\cdot\mathbf{u})$$

Esto es, (1) se cumple con **z** ortogonal a **u** si, y sólo si, $\alpha = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$ y $\hat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$. El vector $\hat{\mathbf{y}}$ es la **proyección ortogonal de y sobre u**, y el vector **z** es la **componente de y ortogonal a u**.

Si c es cualquier escalar diferente de cero y se reemplaza **u** por $c\mathbf{u}$ en la definición de $\hat{\mathbf{y}}$, entonces la proyección ortogonal de **y** sobre $c\mathbf{u}$ es exactamente la misma proyección ortogonal de **y** sobre **u** (ejercicio 31). De aquí que esta proyección esté determinada por el *subespacio* L generado mediante **u** (la línea que pasa por **u** y **0**). Algunas veces $\hat{\mathbf{y}}$ se denota con $\text{proy}_L\,\mathbf{y}$ y se le llama **proyección ortogonal de y sobre** L. Esto es,

$$\boxed{\hat{\mathbf{y}} = \text{proy}_L\,\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}} \tag{2}$$



EJEMPLO 3 Sean $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Encuentre la proyección ortogonal de **y** sobre **u**. Luego escriba **y** como la suma de dos vectores ortogonales, uno en Gen{**u**} y otro ortogonal a **u**.

Solución Calcule

$$\mathbf{y}\cdot\mathbf{u} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}\cdot\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$$

$$\mathbf{u}\cdot\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}\cdot\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

La proyección ortogonal de **y** sobre **u** es

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{40}{20}\mathbf{u} = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

y la componente de **y** ortogonal a **u** es

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

La suma de estos dos vectores es **y**. Es decir,

$$\underset{\substack{\uparrow \\ \mathbf{y}}}{\begin{bmatrix} 7 \\ 6 \end{bmatrix}} = \underset{\substack{\uparrow \\ \hat{\mathbf{y}}}}{\begin{bmatrix} 8 \\ 4 \end{bmatrix}} + \underset{\substack{\uparrow \\ (\mathbf{y}-\hat{\mathbf{y}})}}{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}$$

Esta descomposición de **y** se ilustra en la figura 3. *Nota:* Si los cálculos anteriores son correctos, entonces $\{\hat{\mathbf{y}}, \mathbf{y} - \hat{\mathbf{y}}\}$ será un conjunto ortogonal. Como comprobación, calcule

$$\hat{\mathbf{y}}\cdot(\mathbf{y} - \hat{\mathbf{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix}\cdot\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$

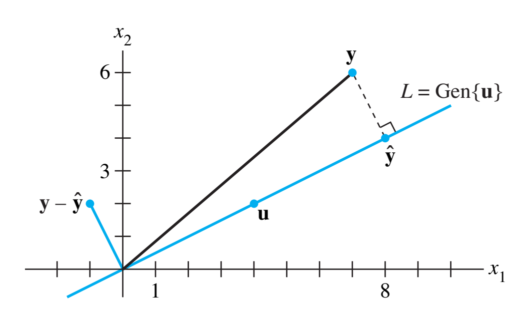


FIGURA 3 La proyección ortogonal de **y** sobre una línea L que pasa por el origen.

Dado que en la figura 3 el segmento de línea entre **y** y $\hat{\mathbf{y}}$ es perpendicular a L, gracias a la estructuración de $\hat{\mathbf{y}}$, el punto identificado con $\hat{\mathbf{y}}$ es el punto de L más cercano a **y**. (Es posible demostrar lo anterior mediante geometría. Se supondrá esto ahora para $\mathbb{R}^2$ y se probará para $\mathbb{R}^n$ en la sección 6.3.)

EJEMPLO 4 Encuentre la distancia de **y** a L en la figura 3.

Solución La distancia de $\mathbf{y}$ a L es la longitud del segmento de línea perpendicular que va desde $\mathbf{y}$ hasta la proyección ortogonal $\hat{\mathbf{y}}$. Esta longitud es igual a la longitud de $\mathbf{y} - \hat{\mathbf{y}}$. Entonces la distancia es

$$\|\mathbf{y} - \hat{\mathbf{y}}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$$

Una interpretación geométrica del teorema 5

La fórmula para la proyección ortogonal $\hat{\mathbf{y}}$ en (2) tiene la misma apariencia que cada uno de los términos del teorema 5. Así, el teorema 5 descompone un vector $\mathbf{y}$ en una suma de proyecciones ortogonales sobre subespacios unidimensionales.

Es fácil visualizar el caso en que $W = \mathbb{R}^2 = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$, siendo $\mathbf{u}_1$ y $\mathbf{u}_2$ ortogonales. Cualquier $\mathbf{y}$ en $\mathbb{R}^2$ puede escribirse en la forma

$$\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 \tag{3}$$

El primer término que aparece en (3) es la proyección de $\mathbf{y}$ sobre el subespacio generado por $\mathbf{u}_1$ (la línea que pasa por $\mathbf{u}_1$ y el origen), y el segundo término es la proyección de $\mathbf{y}$ sobre el subespacio generado por $\mathbf{u}_2$. De manera que (3) expresa a $\mathbf{y}$ como la suma de sus proyecciones sobre los ejes (ortogonales) determinados por $\mathbf{u}_1$ y $\mathbf{u}_2$. Vea la figura 4.

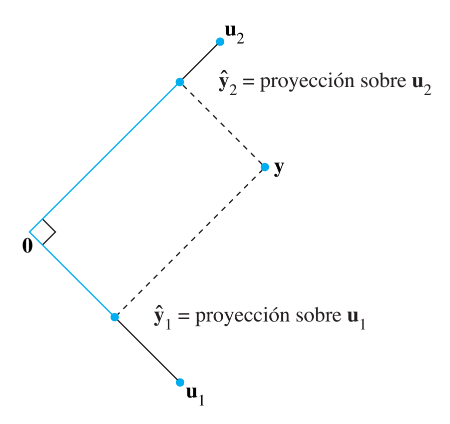


FIGURA 4 Un vector descompuesto en la suma de dos proyecciones.

El teorema 5 descompone cada $\mathbf{y}$ de $\text{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ en la suma de p proyecciones sobre los subespacios unidimensionales que son mutuamente ortogonales.

Descomposición de una fuerza en fuerzas componentes

En física, la descomposición de la figura 4 puede ocurrir cuando algún tipo de fuerza es aplicado a un objeto. Al seleccionarse un sistema de coordenadas apropiado, la fuerza

se representa mediante un vector **y** en $\mathbb{R}^2$ o $\mathbb{R}^3$. Es común que en el problema intervenga alguna dirección de interés particular, la cual se representa con otro vector **u**. Por ejemplo, si el objeto se mueve en línea recta cuando se aplica la fuerza, el vector **u** podría apuntar en la dirección del movimiento, como en la figura 5. Un paso clave del problema consiste en descomponer la fuerza en una componente que vaya en dirección de **u** y otra componente que sea ortogonal a **u**. Los cálculos serían análogos a los efectuados antes en el ejemplo 3.

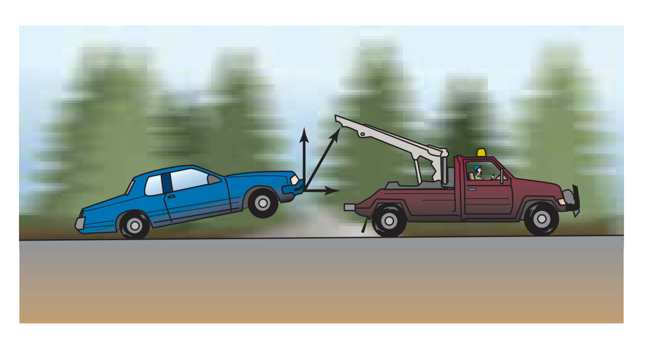


FIGURA 5

Conjuntos ortonormales

Un conjunto $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es un **conjunto ortonormal** si es un conjunto ortogonal de vectores unitarios. Si W es el subespacio generado por un conjunto de este tipo, entonces $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una **base ortonormal** para W, puesto que el conjunto es, de manera automática, linealmente independiente, según el teorema 4.

El ejemplo más sencillo de un conjunto ortonormal es la base estándar $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ para $\mathbb{R}^n$. Cualquier subconjunto no vacío de $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ también es ortonormal. A continuación se presenta un ejemplo más complicado.

EJEMPLO 5 Muestre que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortonormal de $\mathbb{R}^3$, donde

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

Solución Calcule

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_2 &= -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0 \\ \mathbf{v}_1 \cdot \mathbf{v}_3 &= -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0 \\ \mathbf{v}_2 \cdot \mathbf{v}_3 &= 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0 \end{aligned}$$

Entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal. También,

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_1 &= 9/11 + 1/11 + 1/11 = 1 \\ \mathbf{v}_2 \cdot \mathbf{v}_2 &= 1/6 + 4/6 + 1/6 = 1 \\ \mathbf{v}_3 \cdot \mathbf{v}_3 &= 1/66 + 16/66 + 49/66 = 1 \end{aligned}$$

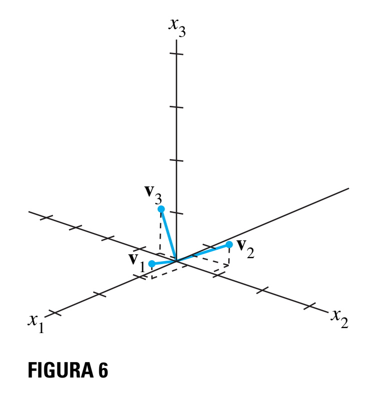


FIGURA 6

lo cual muestra que $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ son vectores unitarios. Entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortonormal. Como el conjunto es linealmente independiente, sus tres vectores forman una base para $\mathbb{R}^3$. Vea la figura 6. ■

Cuando los vectores de un conjunto ortogonal se *normalizan* para tener una longitud unitaria, los nuevos vectores siguen siendo ortogonales, y, por lo tanto, el nuevo conjunto será un conjunto ortonormal. Vea el ejercicio 32. Resulta fácil comprobar que los vectores de la figura 6 (ejemplo 5) son simplemente los vectores unitarios en las direcciones de los vectores de la figura 1 (ejemplo 1).

Las matrices cuyas columnas forman un conjunto ortonormal son importantes en aplicaciones y en algoritmos de computadora para cálculos con matrices. Sus propiedades principales se presentan en los teoremas 6 y 7.

TEOREMA 6 Una matriz U de $m \times n$ tiene columnas ortonormales si, y sólo si, $U^TU = I$.

DEMOSTRACIÓN Para simplificar la notación, se supone que U tiene sólo tres columnas, y cada columna un vector en $\mathbb{R}^m$. La demostración del caso general es esencialmente la misma. Sea $U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3]$ y calcule

$$U^TU = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix} \tag{4}$$

Las entradas de la matriz situada a la derecha son productos interiores, usando notación transpuesta. Las columnas de U son ortogonales si, y sólo si,

$$\mathbf{u}_1^T\mathbf{u}_2 = \mathbf{u}_2^T\mathbf{u}_1 = 0, \quad \mathbf{u}_1^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_1 = 0, \quad \mathbf{u}_2^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_2 = 0 \tag{5}$$

Las columnas de U son todas de longitud unitaria si, y sólo si,

$$\mathbf{u}_1^T\mathbf{u}_1 = 1, \quad \mathbf{u}_2^T\mathbf{u}_2 = 1, \quad \mathbf{u}_3^T\mathbf{u}_3 = 1 \tag{6}$$

El teorema se deriva inmediatamente de las ecuaciones (4) a (6). ■

TEOREMA 7 Sea U una matriz de $m \times n$ con columnas ortonormales, y sean $\mathbf{x}$ y $\mathbf{y}$ vectores en $\mathbb{R}^n$. Entonces

a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$
b. $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$
c. $(U\mathbf{x})\cdot(U\mathbf{y}) = 0$ si, y sólo si, $\mathbf{x}\cdot\mathbf{y} = 0$

Las propiedades (a) y (c) postulan que la función lineal $\mathbf{x} \mapsto U\mathbf{x}$ conserva longitudes y ortogonalidad. Estas propiedades son cruciales para muchos algoritmos de computadora. Para la demostración del teorema 7, vea el ejercicio 25.

EJEMPLO 6 Sean $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Observe que U tiene columnas ortonormales y

$$U^TU = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Verifique si $\|U\mathbf{x}\| = \|\mathbf{x}\|$.

Solución

$$U\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|U\mathbf{x}\| = \sqrt{9+1+1} = \sqrt{11}$$
$$\|\mathbf{x}\| = \sqrt{2+9} = \sqrt{11}$$

Los teoremas 6 y 7 resultan particularmente útiles cuando se aplican a matrices *cuadradas*. Una **matriz ortogonal** es una matriz U cuadrada invertible tal que $U^{-1} = U^T$. De acuerdo con el teorema 6, una matriz de este tipo tiene columnas ortonormales.[1] Resulta fácil advertir que cualquier matriz *cuadrada* con columnas ortonormales es una matriz ortogonal. De manera sorpresiva, tal matriz también debe tener *filas* ortonormales. Vea los ejercicios 27 y 28. En el capítulo 7, las matrices ortogonales se usarán ampliamente.

EJEMPLO 7 La matriz

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

es una matriz ortogonal porque es cuadrada y sus columnas son ortonormales, según el ejemplo 5. Verifique si también las filas son ortonormales.

PROBLEMAS DE PRÁCTICA

1. Sean $\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$. Muestre que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortonormal para $\mathbb{R}^2$.
2. Sean $\mathbf{y}$ y L como en el ejemplo 3 y la figura 3. Determine la proyección ortogonal $\hat{\mathbf{y}}$ de $\mathbf{y}$ sobre L usando $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ en lugar de la $\mathbf{u}$ del ejemplo 3.
3. Sean U y $\mathbf{x}$ como en el ejemplo 6, y sea $\mathbf{y} = \begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix}$. Verifique que $U\mathbf{x} \cdot U\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$.

[1] Un mejor nombre podría ser *matriz ortonormal*; incluso es posible encontrarse con este término en algunos textos de estadística. Sin embargo, en álgebra lineal, el término estándar es *matriz ortogonal*.

6.2 EJERCICIOS

En los ejercicios 1 a 6, determine cuáles conjuntos de vectores son ortogonales.

1. $\begin{bmatrix} -1 \\ 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -7 \end{bmatrix}$

2. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -7 \\ -1 \end{bmatrix}, \begin{bmatrix} -6 \\ -3 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$

4. $\begin{bmatrix} 2 \\ -5 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix}$

5. $\begin{bmatrix} 3 \\ -2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ 7 \\ 0 \end{bmatrix}$

6. $\begin{bmatrix} 5 \\ -4 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ -3 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 5 \\ -1 \end{bmatrix}$

En los ejercicios 7 a 10, muestre que $\{\mathbf{u}_1, \mathbf{u}_2\}$ o $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es una base ortogonal para $\mathbb{R}^2$ o $\mathbb{R}^3$, respectivamente. Después exprese $\mathbf{x}$ como una combinación lineal de las $\mathbf{u}$.

7. $\mathbf{u}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$

8. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 6 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$

9. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$, y $\mathbf{x} = \begin{bmatrix} 8 \\ -4 \\ -3 \end{bmatrix}$

10. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$, y $\mathbf{x} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$

11. Determine la proyección ortogonal de $\begin{bmatrix} 1 \\ 7 \end{bmatrix}$ sobre la línea que pasa por $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ y el origen.

12. Determine la proyección ortogonal de $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ sobre la línea que pasa por $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ y el origen.

13. Sea $\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$. Escriba $\mathbf{y}$ como la suma de dos vectores ortogonales, uno en Gen$\{\mathbf{u}\}$ y otro ortogonal a $\mathbf{u}$.

14. Sea $\mathbf{y} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Escriba $\mathbf{y}$ como la suma de un vector en Gen$\{\mathbf{u}\}$ y un vector ortogonal a $\mathbf{u}$.

15. Sea $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$. Determine la distancia de $\mathbf{y}$ a la línea que pasa por $\mathbf{u}$ y el origen.

16. Sea $\mathbf{y} = \begin{bmatrix} -3 \\ 9 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Determine la distancia de $\mathbf{y}$ a la línea que pasa por $\mathbf{u}$ y el origen.

En los ejercicios 17 a 22, determine cuáles conjuntos de vectores son ortonormales. Si un conjunto es solamente ortogonal, normalice los vectores para producir un conjunto ortonormal.

17. $\begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 0 \\ 1)2 \end{bmatrix}$

18. $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$

19. $\begin{bmatrix} -.6 \\ .8 \end{bmatrix}, \begin{bmatrix} .8 \\ .6 \end{bmatrix}$

20. $\begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 1/3 \\ 2/3 \\ 0 \end{bmatrix}$

21. $\begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{20} \\ 3/\sqrt{20} \end{bmatrix}, \begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{20} \\ -1/\sqrt{20} \end{bmatrix}, \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

22. $\begin{bmatrix} 1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$

En los ejercicios 23 y 24, todos los vectores están en $\mathbb{R}^n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

23. a. No todo conjunto linealmente independiente en $\mathbb{R}^n$ es un conjunto ortogonal.

b. Si $\mathbf{y}$ es una combinación lineal de vectores diferentes de cero a partir de un conjunto ortogonal, entonces los pesos de la combinación lineal pueden calcularse sin aplicar operaciones por fila sobre una matriz.

c. Si los vectores de un conjunto ortogonal de vectores diferentes de cero se normalizan, entonces puede ser que algunos de los nuevos vectores no sean ortogonales.

d. Una matriz con columnas ortonormales es una matriz ortogonal.

e. Si L es una línea que pasa por $\mathbf{0}$ y si $\hat{\mathbf{y}}$ es la proyección ortogonal de $\mathbf{y}$ sobre L, entonces $\|\hat{\mathbf{y}}\|$ proporciona la distancia de $\mathbf{y}$ a L.

24. a. No todo conjunto ortogonal en $\mathbb{R}^n$ es linealmente independiente.

b. Si un conjunto $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ tiene la propiedad de que $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ siempre que $i \neq j$, entonces S es un conjunto ortonormal.

c. Si las columnas de una matriz A de $m \times n$ son ortonormales, entonces la función lineal $\mathbf{x} \mapsto A\mathbf{x}$ conserva las longitudes.

d. La proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{v}$ es la misma que la proyección ortogonal de $\mathbf{y}$ sobre $c\mathbf{v}$ siempre que $c \neq 0$.

e. Una matriz ortogonal es invertible.

25. Demuestre el teorema 7. [*Sugerencia:* Para (a), calcule $\|U\mathbf{x}\|^2$, o demuestre primero (b).]

26. Suponga que W es un subespacio de $\mathbb{R}^n$ generado por n vectores ortogonales diferentes de cero. Explique por qué $W = \mathbb{R}^n$.

27. Sea U una matriz cuadrada con columnas ortonormales. Explique por qué U es invertible. (Mencione los teoremas que utilice.)

28. Sea U una matriz ortogonal de $n \times n$. Demuestre que las filas de U forman una base ortonormal de $\mathbb{R}^n$.

29. Sean U y V matrices ortogonales. Explique por qué UV es una matriz ortogonal. [Es decir, explique por qué UV es invertible y su inverso es $(UV)^T$.]

30. Sea U una matriz ortogonal, y construya V al intercambiar algunas de las columnas de U. Explique por qué V es ortogonal.

31. Demuestre que la proyección ortogonal de un vector $\mathbf{y}$ sobre una línea L que pasa por el origen en $\mathbb{R}^2$ no depende de la elección del $\mathbf{u}$ en L diferente de cero usado en la fórmula para $\hat{\mathbf{y}}$. Para hacer esto, suponga que $\mathbf{y}$ y $\mathbf{u}$ están dados y que $\hat{\mathbf{y}}$ se ha calculado mediante la fórmula (2) de esta sección. Reemplace $\mathbf{u}$ en esa fórmula por $c\mathbf{u}$, donde c es un escalar diferente de cero no especificado. Demuestre que la nueva fórmula proporciona el mismo $\hat{\mathbf{y}}$.

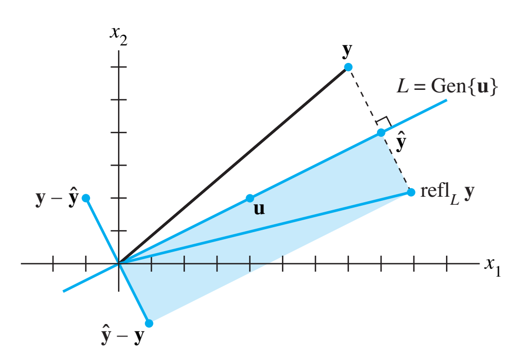
32. Sea $\{\mathbf{v}_1, \mathbf{v}_2\}$ un conjunto ortogonal de vectores diferentes de cero, y c_1 y c_2 cualesquiera escalares diferentes de cero. Muestre que $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2\}$ también es un conjunto ortogonal. Como la ortogonalidad de un conjunto se define en términos de pares de vectores, esto demuestra que si los vectores de un conjunto ortogonal se normalizan, el conjunto nuevo seguirá siendo ortogonal.

33. Dado $\mathbf{u} \neq \mathbf{0}$ en $\mathbb{R}^n$, sea $L = \text{Gen}\{\mathbf{u}\}$. Muestre que la función $\mathbf{x} \mapsto \text{proj}_L\mathbf{x}$ es una transformación lineal.

34. Dado que $\mathbf{u} \neq \mathbf{0}$ en $\mathbb{R}^n$, sea $L = \text{Gen}\{\mathbf{u}\}$. Para $\mathbf{y}$ en $\mathbb{R}^n$, la **reflexión de y en** L es el punto $\text{refl}_L\mathbf{y}$ definido mediante

$$\text{refl}_L\mathbf{y} = 2 \cdot \text{proy}_L\mathbf{y} - \mathbf{y}.$$

Vea la figura, la cual muestra que $\text{refl}_L\mathbf{y}$ es la suma de $\hat{\mathbf{y}} = \text{proy}_L\mathbf{y}$ y $\hat{\mathbf{y}} - \mathbf{y}$. Muestre que la función $\mathbf{y} \mapsto \text{refl}_L\mathbf{y}$ es una transformación lineal.



La reflexión de $\mathbf{y}$ en una línea que pasa por el origen.

35. **[M]** Mediante un cálculo matricial adecuado, muestre que las columnas de la matriz A son ortogonales. Especifique el cálculo que utilice.

$$A = \begin{bmatrix} -6 & -3 & 6 & 1 \\ -1 & 2 & 1 & -6 \\ 3 & 6 & 3 & -2 \\ 6 & -3 & 6 & -1 \\ 2 & -1 & 2 & 3 \\ -3 & 6 & 3 & 2 \\ -2 & -1 & 2 & -3 \\ 1 & 2 & 1 & 6 \end{bmatrix}$$

36. **[M]** En los incisos (a) a (d), sea U la matriz formada al normalizar cada columna de la matriz A en el ejercicio 35.

a. Calcule U^TU y UU^T. ¿En qué difieren?

b. Genere un vector $\mathbf{y}$ aleatorio en $\mathbb{R}^8$ y calcule $\mathbf{p} = UU^T\mathbf{y}$ y $\mathbf{z} = \mathbf{y} - \mathbf{p}$. Explique por qué $\mathbf{p}$ está en Col A. Verifique si $\mathbf{z}$ es ortogonal a $\mathbf{p}$.

c. Compruebe que $\mathbf{z}$ es ortogonal a cada columna de U.

d. Observe que $\mathbf{y} = \mathbf{p} + \mathbf{z}$, con $\mathbf{p}$ en Col A. Explique por qué $\mathbf{z}$ está en $(\text{Col } A)^{\perp}$. (La importancia de esta descomposición de $\mathbf{y}$ se explicará en la siguiente sección.)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Los vectores son ortogonales porque

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = -2/5 + 2/5 = 0$$

Son vectores unitarios porque

$$\|\mathbf{u}_1\|^2 = (-1/\sqrt{5})^2 + (2/\sqrt{5})^2 = 1/5 + 4/5 = 1$$
$$\|\mathbf{u}_2\|^2 = (2/\sqrt{5})^2 + (1/\sqrt{5})^2 = 4/5 + 1/5 = 1$$

En particular, el conjunto $\{\mathbf{u}_1, \mathbf{u}_2\}$ es linealmente independiente, y, por lo tanto, una base para $\mathbb{R}^2$ puesto que hay dos vectores en el conjunto.

2. Cuando $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{20}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

Éste es el mismo $\hat{\mathbf{y}}$ que se encontró en el ejemplo 3. La proyección ortogonal parece depender del $\mathbf{u}$ que se elija en la línea. Vea el ejercicio 31.

3. $U\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}\begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix}$

SG Dominio de bases ortogonales 6 a 4 (Mastering: Orthogonal Bases 6-4)

También, a partir del ejemplo 6, $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$ y $U\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$. Por lo tanto,

$$U\mathbf{x}\cdot U\mathbf{y} = 3 + 7 + 2 = 12, \quad \text{y} \quad \mathbf{x}\cdot\mathbf{y} = -6 + 18 = 12$$

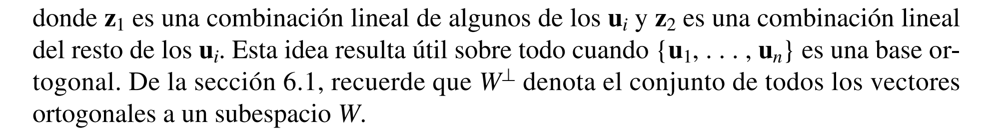
6.3 PROYECCIONES ORTOGONALES

Las proyecciones ortogonales de un punto en $\mathbb{R}^2$ sobre una línea que pasa por el origen tienen un importante análogo en $\mathbb{R}^n$. Dado un vector $\mathbf{y}$ y un subespacio W en $\mathbb{R}^n$, existe un vector $\hat{\mathbf{y}}$ en W tal que (1) $\hat{\mathbf{y}}$ es el único vector de W para el cual $\hat{\mathbf{y}} - \mathbf{y}$ es ortogonal a W, y (2) $\hat{\mathbf{y}}$ es el vector único de W más cercano a $\mathbf{y}$. Vea la figura 1. Estas dos propiedades de $\hat{\mathbf{y}}$ proporcionan la clave para encontrar soluciones a sistemas lineales mediante mínimos cuadrados, y se mencionaron en el ejemplo introductorio de este capítulo. La historia completa será conocida en la sección 6.5.

FIGURA 1

Como preparación para el primer teorema, se observa que siempre que un vector $\mathbf{y}$ se escriba como una combinación de vectores $\mathbf{u}_1, \ldots, \mathbf{u}_n$ en una base de $\mathbb{R}^n$, los términos de la suma para $\mathbf{y}$ podrán agruparse en dos partes de manera que $\mathbf{y}$ pueda escribirse como

$$\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$$


donde $\mathbf{z}_1$ es una combinación lineal de algunos de los $\mathbf{u}_i$ y $\mathbf{z}_2$ es una combinación lineal del resto de los $\mathbf{u}_i$. Esta idea resulta útil sobre todo cuando $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ es una base ortogonal. De la sección 6.1, recuerde que $W^\perp$ denota el conjunto de todos los vectores ortogonales a un subespacio W.

EJEMPLO 1 Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_5\}$ una base ortogonal para $\mathbb{R}^5$ y sea

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_5\mathbf{u}_5$$

Considere el subespacio $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$, y escriba $\mathbf{y}$ como la suma de un vector $\mathbf{z}_1$ en W y un vector $\mathbf{z}_2$ en $W^\perp$.

Solución Escriba

$$\mathbf{y} = \underbrace{c_1\mathbf{u}_1 + c_2\mathbf{u}_2}_{\mathbf{z}_1} + \underbrace{c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5}_{\mathbf{z}_2}$$

donde $\mathbf{z}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ está en Gen $\{\mathbf{u}_1, \mathbf{u}_2\}$

y $\mathbf{z}_2 = c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5$ está en Gen $\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$.

Para mostrar que $\mathbf{z}_2$ está en $W^{\perp}$, basta con probar que $\mathbf{z}_2$ es ortogonal a los vectores de la base $\{\mathbf{u}_1, \mathbf{u}_2\}$ para W. (Vea la sección 6.1.) Utilice las propiedades del producto interior para calcular

$$\begin{aligned}\mathbf{z}_2 \cdot \mathbf{u}_1 &= (c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5) \cdot \mathbf{u}_1 \\ &= c_3\mathbf{u}_3 \cdot \mathbf{u}_1 + c_4\mathbf{u}_4 \cdot \mathbf{u}_1 + c_5\mathbf{u}_5 \cdot \mathbf{u}_1 \\ &= 0\end{aligned}$$

porque $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_3$, $\mathbf{u}_4$ y $\mathbf{u}_5$. Un cálculo semejante muestra que $\mathbf{z}_2 \cdot \mathbf{u}_2 = 0$. Entonces $\mathbf{z}_2$ está en $W^{\perp}$.

El teorema siguiente muestra que la descomposición $\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$ del ejemplo 1 puede calcularse sin tener una base ortogonal para $\mathbb{R}^n$. Basta con tener una base ortogonal sólo para W.

TEOREMA 8

El teorema de la descomposición ortogonal

Sea W un subespacio de $\mathbb{R}^n$. Entonces toda $\mathbf{y}$ en $\mathbb{R}^n$ puede escribirse únicamente en la forma

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

donde $\hat{\mathbf{y}}$ está en W y $\mathbf{z}$ en $W^{\perp}$. De hecho, si $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es cualquier base ortogonal de W, entonces

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p \tag{2}$$

y $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

El vector $\hat{\mathbf{y}}$ de (1) es la **proyección ortogonal de y sobre** W y a menudo se escribe como $\text{proy}_W\, \mathbf{y}$. Vea la figura 2. Cuando W es un subespacio unidimensional, la fórmula para $\hat{\mathbf{y}}$ corresponde a la fórmula dada en la sección 6.2.

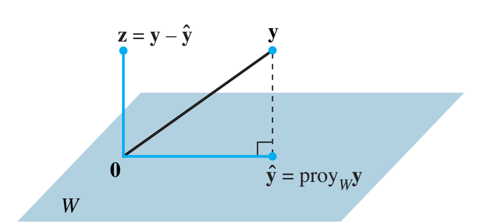


FIGURA 2 La proyección ortogonal de **y** sobre W.

DEMOSTRACIÓN Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ una base ortogonal para W y defina $\hat{\mathbf{y}}$ mediante (2).[1] Entonces $\hat{\mathbf{y}}$ está en W porque $\hat{\mathbf{y}}$ es una combinación lineal de la base $\mathbf{u}_1, \ldots, \mathbf{u}_p$. Sea $< \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. Como $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_2, \ldots, \mathbf{u}_p$ se deduce a partir de (2) que

$$\begin{aligned} \mathbf{z}\cdot\mathbf{u}_1 = (\mathbf{y} - \hat{\mathbf{y}})\cdot\mathbf{u}_1 &= \mathbf{y}\cdot\mathbf{u}_1 - \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1\cdot\mathbf{u}_1 - 0 - \cdots - 0 \\ &= \mathbf{y}\cdot\mathbf{u}_1 - \mathbf{y}\cdot\mathbf{u}_1 = 0 \end{aligned}$$

Entonces $\mathbf{z}$ es ortogonal a $\mathbf{u}_1$. De manera semejante, $\mathbf{z}$ es ortogonal a cada $\mathbf{u}_j$ en la base para W. Por lo tanto, $\mathbf{z}$ es ortogonal a todo vector de W. Es decir, $\mathbf{z}$ está en $W^\perp$.

Para mostrar que la descomposición en (1) es única, suponga que $\mathbf{y}$ también puede escribirse como $\mathbf{y} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$, con $\hat{\mathbf{y}}_1$ en W y $\mathbf{z}_1$ en $W^\perp$. Entonces $\hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$ (puesto que ambos lados son iguales a $\mathbf{y}$), y por ende

$$\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}$$

Esta igualdad muestra que el vector $\mathbf{v} = \hat{\mathbf{y}} - \hat{\mathbf{y}}_1$ está en W y en $W^\perp$ (porque tanto $\mathbf{z}_1$ como $\mathbf{z}$ están en $W^\perp$ y $W^\perp$ es un subespacio). Por lo tanto, $\mathbf{v}\cdot\mathbf{v} = 0$, lo cual muestra que $\mathbf{v} = \mathbf{0}$. Esto demuestra que $\hat{\mathbf{y}} = \hat{\mathbf{y}}_1$ y que $\mathbf{z}_1 = \mathbf{z}$. ■

La unicidad de la descomposición (1) muestra que la proyección ortogonal $\hat{\mathbf{y}}$ depende sólo de W y no de la base específica utilizada en (2).

EJEMPLO 2 Sean $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, y $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortogonal para $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$. Escriba $\mathbf{y}$ como la suma de un vector en W y un vector ortogonal a W.

Solución La proyección ortogonal de $\mathbf{y}$ sobre W es

$$\begin{aligned} \hat{\mathbf{y}} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 \\ &= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} \end{aligned}$$

También

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

El teorema 8 asegura que $\mathbf{y} - \hat{\mathbf{y}}$ está en $W^\perp$. Sin embargo, para comprobar los cálculos, es buena idea verificar que $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal tanto a $\mathbf{u}_1$ como a $\mathbf{u}_2$ y, por lo tanto, a

[1] Puede suponerse que W no es el subespacio cero, porque de otra manera $W^\perp = \mathbb{R}^n$ y (1) es simplemente $\mathbf{y} = \mathbf{0} + \mathbf{y}$. En la siguiente sección se demostrará que cualquier subespacio de $\mathbb{R}^n$ diferente de cero tiene una base ortogonal.

todo W. La descomposición deseada de $\mathbf{y}$ es

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

Una interpretación geométrica de la proyección ortogonal

Cuando W es un subespacio unidimensional, la fórmula (2) para $\text{proy}_W\,\mathbf{y}$ sólo contiene un término. Entonces, cuando $\dim W > 1$, cada término de (2) es él mismo una proyección ortogonal de $\mathbf{y}$ sobre un subespacio unidimensional generado por uno de los $\mathbf{u}$ de la base para W. En la figura 3 se ilustra esto cuando W es un subespacio de $\mathbb{R}^3$ generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Aquí $\hat{\mathbf{y}}_1$ y $\hat{\mathbf{y}}_2$ denotan las proyecciones de $\mathbf{y}$ sobre las líneas generadas por $\mathbf{u}_1$ y $\mathbf{u}_2$, respectivamente. La proyección ortogonal $\hat{\mathbf{y}}$ de $\mathbf{y}$ sobre W es la suma de las proyecciones de $\mathbf{y}$ sobre subespacios unidimensionales que son ortogonales entre sí. El vector $\hat{\mathbf{y}}$ de la figura 3 corresponde al vector $\mathbf{y}$ de la figura 4 mostrado en la sección 6.2, porque ahora es $\hat{\mathbf{y}}$ el que está en W.

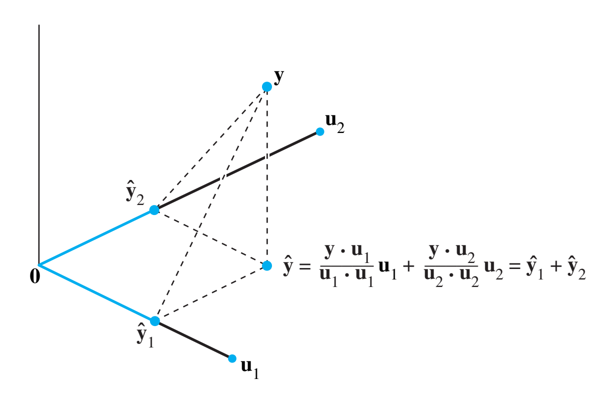


FIGURA 3 La proyección ortogonal de $\mathbf{y}$ es la suma de sus proyecciones sobre subespacios unidimensionales que son mutuamente ortogonales.

Propiedades de las proyecciones ortogonales

Si $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una base ortogonal para W, y si sucede que $\mathbf{y}$ está en W, entonces la fórmula para $\text{proy}_W\,\mathbf{y}$ es exactamente la misma que la proporcionada para la representación de $\mathbf{y}$ en el teorema 5 de la sección 6.2. En este caso, $\text{proy}_W\,\mathbf{y} = \mathbf{y}$.

Si $\mathbf{y}$ está en $W = \text{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, entonces $\text{proy}_W\,\mathbf{y} = \mathbf{y}$.

Este hecho también se deriva del teorema siguiente.

TEOREMA 9

El teorema de la mejor aproximación

Sean W un subespacio de $\mathbb{R}^n$, $\mathbf{y}$ cualquier vector en $\mathbb{R}^n$, y $\hat{\mathbf{y}}$ la proyección ortogonal de $\mathbf{y}$ sobre W. Entonces $\hat{\mathbf{y}}$ es el punto de W más cercano a $\mathbf{y}$, en el sentido que

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \tag{3}$$

para todo $\mathbf{v}$ en W distinto de $\hat{\mathbf{y}}$.

El vector $\hat{\mathbf{y}}$ del teorema 9 es **la mejor aproximación a y de los elementos de** W. En secciones posteriores se examinarán problemas en los que un $\mathbf{y}$ específico debe reemplazarse por (o "aproximarse" a) un vector $\mathbf{v}$ de algún subespacio fijo W. La distancia de $\mathbf{y}$ a $\mathbf{v}$, dada por $\|\mathbf{y} - \mathbf{v}\|$, puede considerarse como el "error" de usar $\mathbf{v}$ en lugar de $\mathbf{y}$. El teorema 9 establece que este error se minimiza cuando $\mathbf{v} = \hat{\mathbf{y}}$.

La ecuación (3) conduce a una nueva demostración de que $\hat{\mathbf{y}}$ no depende de la base ortogonal específica usada para calcularla. De utilizarse una base ortogonal de W diferente para estructurar una proyección ortogonal de $\mathbf{y}$, entonces esta proyección también sería el punto más cercano a $\mathbf{y}$ en W, a saber, $\hat{\mathbf{y}}$.

DEMOSTRACIÓN Tome $\mathbf{v}$ en W diferente de $\hat{\mathbf{y}}$. Vea la figura 4. Entonces $\hat{\mathbf{y}} - \mathbf{v}$ está en W. De acuerdo con el teorema de la descomposición ortogonal, $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a W. En particular, $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a $\hat{\mathbf{y}} - \mathbf{v}$ (la cual está en W). Puesto que

$$\mathbf{y} - \mathbf{v} = (\mathbf{y} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \mathbf{v})$$

al aplicar el teorema de Pitágoras se obtiene

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{v}\|^2$$

(Vea el "triángulo rectángulo" sombreado que aparece en la figura 4. Se marca la longitud de cada lado.) Ahora $\|\hat{\mathbf{y}} - \mathbf{v}\|^2 > 0$ porque $\hat{\mathbf{y}} - \mathbf{v} \neq \mathbf{0}$, y así la desigualdad de (3) se deriva inmediatamente. ■

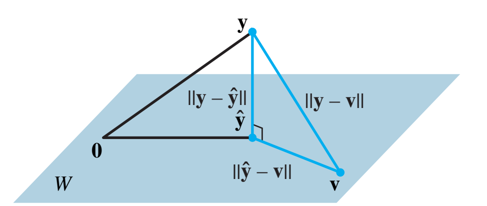


FIGURA 4 La proyección ortogonal de **y** sobre W es el punto más cercano a **y** en W.

EJEMPLO 3 Si $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, y $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$, como en el ejemplo 2, entonces el punto más cercano a $\mathbf{y}$ en W es

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$ ■

EJEMPLO 4 La distancia desde un punto $\mathbf{y}$ en $\mathbb{R}^n$ hasta un subespacio W se define como la distancia desde $\mathbf{y}$ hasta el punto más cercano de W. Encuentre la distancia de $\mathbf{y}$ a $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$, donde

$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Solución De acuerdo con el teorema de la mejor aproximación, la distancia desde $\mathbf{y}$ hasta W es $\|\mathbf{y} - \hat{\mathbf{y}}\|$, donde $\hat{\mathbf{y}} = \text{proy}_W\,\mathbf{y}$. Puesto que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortogonal para W,

$$\hat{\mathbf{y}} = \frac{15}{30}\mathbf{u}_1 + \frac{-21}{6}\mathbf{u}_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = 3^2 + 6^2 = 45$$

La distancia de $\mathbf{y}$ a W es $\sqrt{45} = 3\sqrt{5}$.

El teorema final de la sección muestra cómo la fórmula (2) para $\text{proy}_W\,\mathbf{y}$ se simplifica cuando la base para W es un conjunto ortonormal.

TEOREMA 10

Si $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una base ortonormal para un subespacio W de $\mathbb{R}^n$, entonces

$$\text{proy}_W\,\mathbf{y} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p \tag{4}$$

Si $U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p]$, entonces

$$\text{proy}_W\,\mathbf{y} = UU^T\mathbf{y} \quad \text{para toda } \mathbf{y} \text{ en } \mathbb{R}^n \tag{5}$$

DEMOSTRACIÓN La fórmula (4) es consecuencia inmediata de (2). (4) muestra también que $\text{proy}_W\,\mathbf{y}$ es una combinación lineal de las columnas de U usando los pesos $\mathbf{y}\cdot\mathbf{u}_1$, $\mathbf{y}\cdot\mathbf{u}_2$, $\ldots$, $\mathbf{y}\cdot\mathbf{u}_p$. Los pesos se pueden escribir como $\mathbf{u}_1^T\mathbf{y}, \mathbf{u}_2^T\mathbf{y}, \ldots, \mathbf{u}_p^T\mathbf{y}$, mostrando que son las entradas de $U^T\mathbf{y}$ y justificando (5).

CD La matriz de proyección (The Projection Matrix)

Suponga que U es de $n \times p$ con columnas ortonormales, y sea W el espacio de columnas de U. Entonces

$$\begin{aligned} U^TU\mathbf{x} &= I_p\mathbf{x} = \mathbf{x} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{R}^p \qquad \text{Teorema 6} \\ UU^T\mathbf{y} &= \text{proy}_W\,\mathbf{y} \quad \text{para toda } \mathbf{y} \text{ en } \mathbb{R}^n \qquad \text{Teorema 10} \end{aligned}$$

Si U es una matriz (cuadrada) de $n \times n$ con columnas ortonormales, entonces es una matriz *ortogonal*, el espacio columna W es todo $\mathbb{R}^n$, y $UU^T\mathbf{y} = I\mathbf{y} = \mathbf{y}$ para toda $\mathbf{y}$ en $\mathbb{R}^n$.

Aunque la fórmula (4) es importante para propósitos teóricos, en la práctica requiere usualmente de cálculos con raíces cuadradas de números (en las entradas de $\mathbf{u}_i$). Se recomienda la fórmula (2) para efectuar cálculos a mano.

PROBLEMA DE PRÁCTICA

Sean $\mathbf{u}_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$, y $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$. Utilice el hecho de que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales para calcular $\text{proy}_W\,\mathbf{y}$.

6.3 EJERCICIOS

En los ejercicios 1 y 2, puede suponerse que $\{\mathbf{u}_1, \ldots, \mathbf{u}_4\}$ es una base ortogonal para $\mathbb{R}^4$.

1. $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ -4 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 3 \\ 5 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ -4 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} 5 \\ -3 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 10 \\ -8 \\ 2 \\ 0 \end{bmatrix}$. Escriba $\mathbf{x}$ como la suma de dos vectores, uno en $\text{Gen}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ y el otro en $\text{Gen}\{\mathbf{u}_4\}$.

2. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ -1 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 4 \\ 5 \\ -3 \\ 3 \end{bmatrix}$. Escriba $\mathbf{v}$ como la suma de dos vectores, uno en $\text{Gen}\{\mathbf{u}_1\}$ y el otro en $\text{Gen}\{\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$.

En los ejercicios 3 a 6, compruebe que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es un conjunto ortogonal, y después encuentre la proyección ortogonal de $\mathbf{y}$ sobre $\text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$.

3. $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

4. $\mathbf{y} = \begin{bmatrix} 6 \\ 3 \\ -2 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$

5. $\mathbf{y} = \begin{bmatrix} -1 \\ 2 \\ 6 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$

6. $\mathbf{y} = \begin{bmatrix} 6 \\ 4 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -4 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 7 a 10, sea W el subespacio generado por los $\mathbf{u}$'s, y escriba a $\mathbf{y}$ como la suma de un vector en W y un vector ortogonal a W.

7. $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}$

8. $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$

9. $\mathbf{y} = \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

10. $\mathbf{y} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}$

En los ejercicios 11 y 12, encuentre el punto más cercano a $\mathbf{y}$ en el subespacio W generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

11. $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$

12. $\mathbf{y} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix}$

En los ejercicios 13 y 14, encuentre la mejor aproximación a $\mathbf{z}$ mediante vectores de la forma $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$.

13. $\mathbf{z} = \begin{bmatrix} 3 \\ -7 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ -3 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$

14. $\mathbf{z} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ -3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 5 \\ -2 \\ 4 \\ 2 \end{bmatrix}$

15. Sea $\mathbf{y} = \begin{bmatrix} 5 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$. Encuentre la distancia de $\mathbf{y}$ al plano en $\mathbb{R}^3$ generado por $\mathbf{u}_1$ y $\mathbf{u}_2$.

16. Sean $\mathbf{y}$, $\mathbf{v}_1$ y $\mathbf{v}_2$ como en el ejercicio 12. Encuentre la distancia de $\mathbf{y}$ al subespacio de $\mathbb{R}^4$ generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

17. Sean $\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$, y y $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$.

a. Sea $U = [\mathbf{u}_1 \quad \mathbf{u}_2]$. Calcule U^TU y UU^T.

b. Calcule $\text{proy}_W\, \mathbf{y}$ y $(UU^T)\mathbf{y}$.

18. Sean $\mathbf{y} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$, y $W = \text{Gen}\{\mathbf{u}_1\}$.

a. Sea U la matriz de 2×1 cuya única columna es $\mathbf{u}_1$. Calcule U^TU y UU^T.

b. Calcule $\text{proy}_W\, \mathbf{y}$ y $(UU^T)\mathbf{y}$.

19. Sean $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, y $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Observe que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales pero que $\mathbf{u}_3$ no es ortogonal a $\mathbf{u}_1$ ni a $\mathbf{u}_2$. Es posible demostrar que $\mathbf{u}_3$ no está en el subespacio W generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Utilice este hecho para construir un vector $\mathbf{v}$ diferente de cero en $\mathbb{R}^3$ que sea ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$.

20. Sean $\mathbf{u}_1$ y $\mathbf{u}_2$ como en el ejercicio 19, y sea $\mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Es posible demostrar que $\mathbf{u}_4$ no está en el subespacio W generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Utilice este hecho para construir un vector $\mathbf{v}$ diferente de cero en $\mathbb{R}^3$ que sea ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$.

En los ejercicios 21 y 22, todos los vectores y los subespacios están en $\mathbb{R}^n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

21. a. Si $\mathbf{z}$ es ortogonal a $\mathbf{u}_1$ y a $\mathbf{u}_2$ y si $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$, entonces $\mathbf{z}$ debe estar en $W^\perp$.

b. Para cada $\mathbf{y}$ y cada subespacio W, el vector $\mathbf{y} - \text{proy}_W\, \mathbf{y}$ es ortogonal a W.

c. La proyección ortogonal $\hat{\mathbf{y}}$ de $\mathbf{y}$ sobre un subespacio W puede depender a veces de la base ortogonal para W usada al calcular $\hat{\mathbf{y}}$.

d. Si $\mathbf{y}$ está en un subespacio W, entonces la proyección ortogonal de $\mathbf{y}$ sobre W es $\mathbf{y}$ misma.

e. Si las columnas de una matriz U de $n \times p$ son ortonormales, entonces $UU^T\mathbf{y}$ es la proyección ortogonal de $\mathbf{y}$ sobre el espacio de columnas de U.

22. a. Si W es un subespacio de $\mathbb{R}^n$, y si $\mathbf{v}$ está en W y en $W^\perp$, entonces $\mathbf{v}$ debe ser el vector cero.

b. En el teorema de la descomposición ortogonal, cada término de la fórmula (2) para $\hat{\mathbf{y}}$ es, él mismo, una proyección ortogonal de $\mathbf{y}$ sobre un subespacio de W.

c. Si $\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$, donde $\mathbf{z}_1$ está en un subespacio W y $\mathbf{z}_2$ está en $W^\perp$, entonces $\mathbf{z}_1$ debe ser la proyección ortogonal de $\mathbf{y}$ sobre W.

d. La mejor aproximación a $\mathbf{y}$ con los elementos de un subespacio W está dada por el vector $\mathbf{y} - \text{proy}_W\, \mathbf{y}$.

e. Si una matriz U de $n \times p$ tiene columnas ortonormales, entonces $UU^T\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

23. Sea A una matriz de $m \times n$. Demuestre que todo vector $\mathbf{x}$ en $\mathbb{R}^n$ puede escribirse en la forma $\mathbf{x} = \mathbf{p} + \mathbf{u}$, donde $\mathbf{p}$ está en Fil A y $\mathbf{u}$ en Nul A. También, muestre que si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces hay una $\mathbf{p}$ única en Fil **A** tal que $A\mathbf{p} = \mathbf{b}$.

24. Sea W un subespacio de $\mathbb{R}^n$ con una base ortogonal $\{\mathbf{w}_1, \ldots, \mathbf{w}_p\}$, y sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_q\}$ una base ortogonal para $W^\perp$.

a. Explique porqué $\{\mathbf{w}_1, \ldots, \mathbf{w}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$ es un conjunto ortogonal.

b. Explique por qué el conjunto de la parte (a) genera $\mathbb{R}^n$.

c. Demuestre que $\dim W + \dim W^\perp = n$.

25. **[M]** Sea U la matriz de 8×4 del ejercicio 36 presentado en la sección 6.2. Encuentre el punto más cercano a $\mathbf{y} = (1, 1, 1, 1, 1, 1, 1, 1)$ en Col U. Escriba los comandos o pulsaciones de tecla que utilice para resolver este problema.

26. **[M]** Sea U la matriz del ejercicio 25. Encuentre la distancia de $\mathbf{b} = (1, 1, 1, 1, -1, -1, -1, -1)$ a Col U.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Calcule

$$\text{proy}_W\, \mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \frac{88}{66}\mathbf{u}_1 + \frac{-2}{6}\mathbf{u}_2$$
$$= \frac{4}{3}\begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix} = \mathbf{y}$$

En este caso, $\mathbf{y}$ resulta ser una combinación lineal de $\mathbf{u}_1$ y $\mathbf{u}_2$, así que $\mathbf{y}$ está en W. El punto de W más cercano a $\mathbf{y}$ es $\mathbf{y}$ mismo.

6.4 EL PROCESO GRAM-SCHMIDT

El proceso Gram-Schmidt es un algoritmo sencillo para producir una base ortogonal u ortonormal para cualquier subespacio diferente de cero de $\mathbb{R}^n$. Los primeros dos ejemplos del proceso están enfocados en los cálculos a mano.

EJEMPLO 1 Sea $W = \text{Gen}\{\mathbf{x}_1, \mathbf{x}_2\}$, donde $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ y $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Construya una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$ para W.

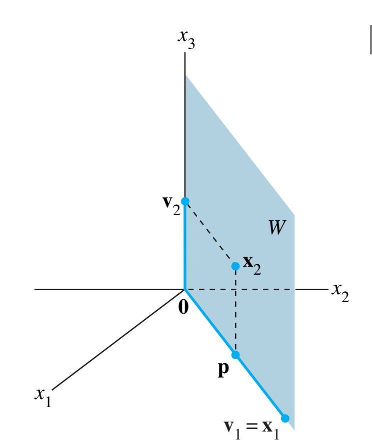


FIGURA 1
Estructuración de una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$.

Solución El subespacio W se muestra en la figura 1, junto con $\mathbf{x}_1$, $\mathbf{x}_2$ y la proyección $\mathbf{p}$ de $\mathbf{x}_2$ sobre $\mathbf{x}_1$. La componente de $\mathbf{x}_2$ ortogonal a $\mathbf{x}_1$ es $\mathbf{x}_2 - \mathbf{p}$, la cual está en W porque se forma a partir de $\mathbf{x}_2$ y de un múltiplo de $\mathbf{x}_1$. Sea $\mathbf{v}_1 = \mathbf{x}_1$ y

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p} = \mathbf{x}_2 - \frac{\mathbf{x}_2\cdot\mathbf{x}_1}{\mathbf{x}_1\cdot\mathbf{x}_1}\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Entonces $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto ortogonal de vectores diferentes de cero en W. Como $\dim W = 2$, el conjunto $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para W.

En el siguiente ejemplo se ilustra el proceso Gram-Schmidt en su totalidad. Estúdielo con cuidado.

EJEMPLO 2 Sean $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Entonces, resulta claro que $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ es linealmente independiente y, por lo tanto, es una base para un subespacio W de $\mathbb{R}^4$. Estructure una base ortogonal para W.

Solución

Paso 1. Sean $\mathbf{v}_1 = \mathbf{x}_1$ y $W_1 = \text{Gen}\{\mathbf{x}_1\} = \text{Gen}\{\mathbf{v}_1\}$.

Paso 2. Sea $\mathbf{v}_2$ el vector producido al restar de $\mathbf{x}_2$ su proyección sobre el subespacio W_1. Esto es, sea

$$\begin{aligned}\mathbf{v}_2 &= \mathbf{x}_2 - \text{proy}_{W_1}\,\mathbf{x}_2 \\ &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \qquad \text{Puesto que } \mathbf{v}_1 = \mathbf{x}_1 \\ &= \begin{bmatrix}0\\1\\1\\1\end{bmatrix} - \frac{3}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}-3/4\\1/4\\1/4\\1/4\end{bmatrix}\end{aligned}$$

Tal como en el ejemplo 1, $\mathbf{v}_2$ es la componente de $\mathbf{x}_2$ ortogonal a $\mathbf{x}_1$, y $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base ortogonal para el subespacio W_2 generado por $\mathbf{x}_1$ y $\mathbf{x}_2$.

Paso 2′ (opcional). Si es apropiado, escale $\mathbf{v}_2$ para simplificar los cálculos posteriores. Como $\mathbf{v}_2$ tiene entradas fraccionarias, es conveniente escalarlo mediante un factor de 4 y reemplazar a $\{\mathbf{v}_1, \mathbf{v}_2\}$ empleando la base ortogonal

$$\mathbf{v}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \qquad \mathbf{v}_2' = \begin{bmatrix}-3\\1\\1\\1\end{bmatrix}$$

Paso 3. Sea $\mathbf{v}_3$ el vector producido al restar de $\mathbf{x}_3$ su proyección sobre el subespacio W_2. Utilice la base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2'\}$ para calcular la proyección sobre W_2:

Proyección de $\mathbf{x}_3$ sobre $\mathbf{v}_1$ ↓ Proyección de $\mathbf{x}_3$ sobre $\mathbf{v}_2'$ ↓

$$\begin{aligned}\text{proy}_{W_2}\,\mathbf{x}_3 &= \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{x}_3 \cdot \mathbf{v}_2'}{\mathbf{v}_2' \cdot \mathbf{v}_2'}\mathbf{v}_2' \\ &= \frac{2}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{2}{12}\begin{bmatrix}-3\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\2/3\\2/3\\2/3\end{bmatrix}\end{aligned}$$

Entonces $\mathbf{v}_3$ es la componente de $\mathbf{x}_3$ ortogonal a W_2, a saber,

$$\mathbf{v}_3 = \mathbf{x}_3 - \text{proy}_{W_2}\,\mathbf{x}_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} - \begin{bmatrix}0\\2/3\\2/3\\2/3\end{bmatrix} = \begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix}$$

En la figura 2, vea un diagrama de esta construcción. Observe que $\mathbf{v}_3$ está en W, porque tanto $\mathbf{x}_3$ como $\text{proy}_{W_2}\,\mathbf{x}_3$ están en W. Entonces $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ es un conjunto ortogonal de vectores diferentes de cero y, por lo tanto, un conjunto linealmente independiente en W. Observe que W es tridimensional, pues fue definido con una base de tres vectores. Por lo tanto, según el teorema de la base presentado en la sección 4.5, $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ es una base ortogonal para W. ■

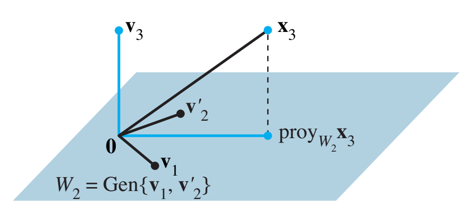


FIGURA 2 Estructuración de $\mathbf{v}_3$ a partir de $\mathbf{x}_3$ y W_2.

La demostración del teorema siguiente pone de manifiesto que esta estrategia sí funciona. No se menciona el escalamiento de los vectores porque eso es algo que se usa únicamente para simplificar los cálculos a mano.

TEOREMA 11

El proceso Gram-Schmidt

Dada una base $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ para un subespacio W de $\mathbb{R}^n$, defina

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es una base ortogonal para W. Además

$$\text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \text{Gen}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \qquad \text{para } 1 \le k \le p \tag{1}$$

DEMOSTRACIÓN Para $i \le k \le p$, sea $W_k = \text{Gen}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$. Haga $\mathbf{v}_1 = \mathbf{x}_1$, de manera que $\text{Gen}\{\mathbf{v}_1\} = \text{Gen}\{\mathbf{x}_1\}$. Suponga que, para alguna $k < p$, se ha construido $\mathbf{v}_1, \ldots, \mathbf{v}_k$ tal que $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es una base ortogonal para W_k. Se define

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - \text{proy}_{W_k}\mathbf{x}_{k+1} \tag{2}$$

De acuerdo con el teorema de la descomposición ortogonal, $\mathbf{v}_{k+1}$ es ortogonal a W_k. Observe que $\text{proy}_{W_k}\mathbf{x}_{k+1}$ está en W_k y, por lo tanto, también está en W_{k+1}. Puesto que $\mathbf{x}_{k+1}$ está en W_{k+1}, también lo está $\mathbf{v}_{k+1}$ (porque W_{k+1} es un subespacio y es cerrado bajo la resta). Más aún, $\mathbf{v}_{k+1} \neq \mathbf{0}$ porque $\mathbf{x}_{k+1}$ no está en $W_k = \text{Gen}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$. De aquí que $\{\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}\}$ sea un conjunto ortogonal de vectores diferentes de cero en el espacio $(k+1)$-dimensional W_{k+1}. Según el teorema de la base presentado en la sección 4.5, este conjunto es una base ortogonal para W_{k+1}. Por lo tanto, $W_{k+1} = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}\}$. Cuando $k + 1 = p$, el proceso se detiene. ■

El teorema 11 muestra que cualquier subespacio W diferente de cero de $\mathbb{R}^n$ tiene una base ortogonal, porque siempre se dispone de una base ordinaria $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ (según el teorema 11 de la sección 4.5), y porque el proceso de Gram-Schmidt depende sólo de

la existencia de proyecciones ortogonales sobre subespacios de W que ya tengan bases ortogonales.

Bases ortonormales

Una base ortonormal se construye fácilmente a partir de una base ortogonal $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$: sólo normalice (es decir, "escale") todas las $\mathbf{v}_k$. Cuando los problemas se resuelven a mano, esto resulta más sencillo que normalizar cada $\mathbf{v}_k$ en cuanto se encuentra (porque se evita escribir raíces cuadradas innecesarias).

EJEMPLO 3 En el ejemplo 1 se estructuró la base ortogonal

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Una base ortonormal es

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{45}}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Factorización QR de matrices

Si una matriz A de $m \times n$ tiene columnas $\mathbf{x}_1, \ldots, \mathbf{x}_n$ linealmente independientes, entonces la aplicación del proceso de Gram-Schmidt (con normalizaciones) a $\mathbf{x}_1, \ldots, \mathbf{x}_n$ equivale a *factorizar* A tal como se describe en el teorema siguiente. Esta factorización es usada ampliamente en los algoritmos de computadora para efectuar diversos cálculos, como la resolución de ecuaciones (que se analiza en la sección 6.5), y la ubicación de valores propios (mencionada en los ejercicios de la sección 5.2).

TEOREMA 12 **La factorización QR**

Si A es una matriz de $m \times n$ con columnas linealmente independientes, entonces puede factorizarse como $A = QR$, donde Q es una matriz de $m \times n$ cuyas columnas forman una base ortonormal para Col A, y R es una matriz invertible triangular superior de $n \times n$ con entradas positivas en su diagonal.

DEMOSTRACIÓN Las columnas de A forman una base $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ para Col A. Empleando la propiedad (1) dada en el teorema 11, estructure una base ortonormal $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ para $W = \text{Col } A$. Esta base puede estructurarse mediante el proceso Gram-Schmidt o de algunas otras formas. Sea

$$Q = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n\,]$$

Para $k = 1, \ldots, n$, $\mathbf{x}_k$ está en Gen$\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$= Gen$\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. Por lo tanto, existen constantes, $r_{1k}, \ldots, r_{kk}$, tales que

$$\mathbf{x}_k = r_{1k}\mathbf{u}_1 + \cdots + r_{kk}\mathbf{u}_k + 0\cdot\mathbf{u}_{k+1} + \cdots + 0\cdot\mathbf{u}_n$$

Puede suponerse que $r_{kk} \geq 0$. (Si $r_{kk} < 0$, multiplique tanto r_{kk} como $\mathbf{u}_k$ por -1.) Esto muestra que $\mathbf{x}_k$ es una combinación lineal de las columnas de Q utilizando como pesos las entradas del vector

$$\mathbf{r}_k = \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Es decir, $\mathbf{x}_k = Q\mathbf{r}_k$ para $k = 1, \ldots, n$. Sea $R = [\mathbf{r}_1 \quad \cdots \quad \mathbf{r}_n]$. Entonces

$$A = [\,\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_n\,] = [\,Q\mathbf{r}_1 \quad \cdots \quad Q\mathbf{r}_n\,] = QR$$

El que R sea invertible se deduce fácilmente del hecho de que las columnas de A son linealmente independientes (ejercicio 19). Como resulta evidente que R es triangular superior, sus entradas diagonales no negativas deben ser positivas. ■

EJEMPLO 4 Encuentre una factorización QR de $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Solución Las columnas de A son los vectores $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ del ejemplo 2. En ese ejemplo se encontró una base ortogonal para Col A = Gen$\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_2' = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Escale $\mathbf{v}_3$ haciendo $\mathbf{v}_3' = 3\mathbf{v}_3$. Luego normalice los tres vectores para obtener $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, y utilice estos vectores como las columnas de Q.

$$Q = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}$$

Por construcción, las primeras k columnas de Q son una base ortonormal de Gen$\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$. A partir de la demostración del teorema 12, $A = QR$ para alguna R. Para encontrar R, observe que $Q^TQ = I$, porque las columnas de Q son ortonormales. Por lo tanto,

$$Q^TA = Q^T(QR) = IR = R$$

y

$$R = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}$$

NOTAS NUMÉRICAS

1. Cuando el proceso Gram-Schmidt se ejecuta en una computadora, puede acumularse un error de redondeo a medida que se calculan los vectores $\mathbf{u}_k$ uno por uno. Para k y j grandes pero no iguales, los productos interiores $\mathbf{u}_j^T\mathbf{u}_k$ podrían no ser lo suficientemente cercanos a cero. Esta pérdida de ortogonalidad puede reducirse sustancialmente al reacomodar el orden de los cálculos.[1] Sin embargo, generalmente se prefiere factorizar QR en vez de aplicar este método Gram-Schmidt modificado ya que se obtiene una base ortonormal más precisa, a pesar de que la factorización requiere casi el doble de operaciones aritméticas.
2. Para producir una factorización QR de una matriz A, generalmente un programa de computadora multiplica A por la izquierda mediante una sucesión de matrices ortogonales hasta que la transforma en una matriz triangular superior. Esta estructuración es análoga a la multiplicación izquierda por matrices elementales que produce una factorización LU de A.

PROBLEMA DE PRÁCTICA

Sea $W = \text{Gen}\{\mathbf{x}_1, \mathbf{x}_2\}$, donde $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{x}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$. Estructure una base ortonormal para W.

6.4 EJERCICIOS

En los ejercicios 1 a 6, el conjunto dado es una base para un subespacio W. Utilice el proceso Gram-Schmidt para producir una base ortogonal de W.

1. $\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 8 \\ 5 \\ -6 \end{bmatrix}$

2. $\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$

4. $\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix}$

[1]Vea *Fundamentals of Matrix Computations*, por David S. Watkins (Nueva York: John Wiley & Sons, 1991), págs. 167-180.

5. $\begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -7 \\ -4 \\ 1 \end{bmatrix}$

6. $\begin{bmatrix} 3 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -5 \\ 9 \\ -9 \\ 3 \end{bmatrix}$

7. Encuentre una base ortonormal para el subespacio generado por los vectores del ejercicio 3.

8. Encuentre una base ortonormal para el subespacio generado por los vectores del ejercicio 4.

Encuentre una base ortogonal para el espacio de columnas de cada matriz de los ejercicios 9 a 12.

9. $\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & 5 & -2 \\ 3 & -7 & 8 \end{bmatrix}$

10. $\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{bmatrix}$

En los ejercicios 13 y 14, las columnas de Q se obtuvieron aplicando el proceso Gram-Schmidt a las columnas de A. Encuentre una matriz triangular superior R tal que $A = QR$. Verifique su trabajo.

13. $A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}, Q = \begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix}$

14. $A = \begin{bmatrix} -2 & 3 \\ 5 & 7 \\ 2 & -2 \\ 4 & 6 \end{bmatrix}, Q = \begin{bmatrix} -2/7 & 5/7 \\ 5/7 & 2/7 \\ 2/7 & -4/7 \\ 4/7 & 2/7 \end{bmatrix}$

15. Encuentre una factorización QR para la matriz del ejercicio 11.

16. Encuentre una factorización QR para la matriz del ejercicio 12.

En los ejercicios 17 y 18, todos los vectores y subespacios están en $\mathbb{R}^n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

17. a. Si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortogonal de W, entonces la multiplicación de $\mathbf{v}_3$ por un escalar c produce una base ortogonal nueva $\{\mathbf{v}_1, \mathbf{v}_2, c\mathbf{v}_3\}$.

b. El proceso Gram-Schmidt produce, a partir de un conjunto linealmente independiente $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$, un conjunto ortogonal $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ con la propiedad de que, para cada k, los vectores generan el mismo subespacio que el generado por $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

c. Si $A = QR$, donde Q tiene columnas ortonormales, entonces $R = Q^TA$.

18. a. Si $W = \text{Gen}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ con $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linealmente independiente, y si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal en W, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base de W.

b. Si $\mathbf{x}$ no está en un subespacio W, entonces $\mathbf{x} - \text{proy}_W\,\mathbf{x}$ no es cero.

c. En una factorización QR, por ejemplo $A = QR$ (cuando A tiene las columnas linealmente independientes), las columnas de Q forman una base ortonormal del espacio de columnas de A.

19. Suponga que $A = QR$, donde Q es de $m \times n$ y R es de $n \times n$. Demuestre que si las columnas de A son linealmente independientes, entonces R debe ser invertible. [*Sugerencia:* Estudie la ecuación $R\mathbf{x} = \mathbf{0}$ y utilice el hecho de que $A = QR$.]

20. Suponga que $A = QR$, donde R es una matriz invertible. Demuestre que A y Q tienen el mismo espacio de columnas. [*Sugerencia:* Dada $\mathbf{y}$ en Col A, muestre que $\mathbf{y} = Q\mathbf{x}$ para alguna $\mathbf{x}$. También, dada $\mathbf{y}$ en Col Q, muestre que $\mathbf{y} = A\mathbf{x}$ para alguna $\mathbf{x}$.]

21. Dada $A = QR$ como en el teorema 12, describa cómo encontrar una matriz (cuadrada) ortogonal Q_1 de $m \times m$ y una matriz de $n \times n$ invertible triangular superior R tales que

$$A = Q_1 \begin{bmatrix} R \\ 0 \end{bmatrix}$$

El comando `qr` de MATLAB proporciona esta factorización QR "completa" cuando rango $A = n$.

22. Si $\mathbf{u}_1, \ldots, \mathbf{u}_p$ es una base ortogonal para un subespacio W de $\mathbb{R}^n$, y $T : \mathbb{R}^n \to \mathbb{R}^n$ se define como $T(\mathbf{x}) = \text{proy}_W\,\mathbf{x}$, muestre que T es una transformación lineal.

23. Suponga que $A = QR$ es una factorización QR de una matriz A de $m \times n$ (con columnas linealmente independientes). Divida A como $[A_1 \quad A_2]$, donde A_1 tiene p columnas. Muestre cómo obtener una factorización QR de A_1, y explique por qué su factorización tiene las propiedades adecuadas.

24. **[M]** Utilice el proceso Gram-Schmidt para producir una base ortogonal del espacio de columnas de

$$A = \begin{bmatrix} -10 & 13 & 7 & -11 \\ 2 & 1 & -5 & 3 \\ -6 & 3 & 13 & -3 \\ 16 & -16 & -2 & 5 \\ 2 & 1 & -5 & -7 \end{bmatrix}$$

25. **[M]** Produzca una factorización QR de la matriz del ejercicio 24.

26. **[M]** Para un programa de matrices, el proceso Gram-Schmidt funciona mejor con vectores ortonormales. Iniciando con $\mathbf{x}_1, \ldots, \mathbf{x}_p$ como en el teorema 11, sea $A = [\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_p]$. Suponga que Q es una matriz de $n \times k$ cuyas columnas forman una base ortonormal del subespacio W_k generado por las pri-

meras k columnas de A. Entonces, para $\mathbf{x}$ en $\mathbb{R}^n$, QQ^T es la proyección ortogonal de $\mathbf{x}$ sobre W_k (teorema 10 de la sección 6.3). Si $\mathbf{x}_{k+1}$ es la siguiente columna de A, entonces la ecuación (2) de la demostración del teorema 11 es

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - Q(Q^T\mathbf{x}_{k+1})$$

(Los paréntesis anteriores reducen el número de operaciones aritméticas.) Sea $\mathbf{u}_{k+1} = \mathbf{v}_{k+1}/\|\mathbf{v}_{k+1}\|$. La nueva Q para el siguiente paso es $[Q \quad \mathbf{u}_{k+1}]$. Use este procedimiento para calcular la factorización QR de la matriz del ejercicio 24. Escriba los comandos o pulsaciones de tecla que utilice.

Proceso Gram-Schmidt y una factorización QR (Gram-Schmidt and a QR Factorization)

Solución al problema de práctica

Sean $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \mathbf{x}_2 - \dfrac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - 0\mathbf{v}_1 = \mathbf{x}_2$. Entonces $\{\mathbf{x}_1, \mathbf{x}_2\}$ ya es ortogonal. Todo lo que se necesita es normalizar los vectores. Sea

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

En lugar de normalizar directamente $\mathbf{v}_2$, normalice $\mathbf{v}'_2 = 3\mathbf{v}_2$:

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}'_2\|}\mathbf{v}'_2 = \frac{1}{\sqrt{1^2 + 1^2 + (-2)^2}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

Entonces $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortonormal de W.

6.5 PROBLEMAS DE MÍNIMOS CUADRADOS

El ejemplo introductorio de este capítulo describe un problema muy grande del tipo $A\mathbf{x} = \mathbf{b}$ que no tenía solución. Los sistemas inconsistentes surgen con frecuencia en las aplicaciones, aunque generalmente no tienen una matriz de coeficientes tan enorme. Cuando se necesita una solución pero no hay ninguna, lo mejor que puede hacerse es encontrar una $\mathbf{x}$ que deje a $A\mathbf{x}$ tan cercana a $\mathbf{b}$ como sea posible.

Piense en $A\mathbf{x}$ como una *aproximación* a $\mathbf{b}$. Cuanto más pequeña sea la distancia entre $\mathbf{b}$ y $A\mathbf{x}$, dada por $\|\mathbf{b} - A\mathbf{x}\|$, mejor será la aproximación. El **problema general de mínimos cuadrados** es encontrar un $\mathbf{x}$ que haga a $\|\mathbf{b} - A\mathbf{x}\|$ tan pequeña como sea posible. El adjetivo "mínimos cuadrados" surge de que $\|\mathbf{b} - A\mathbf{x}\|$ es la raíz cuadrada de una suma de cuadrados.

DEFINICIÓN

Si A es de $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$, una **solución por mínimos cuadrados** de $A\mathbf{x} = \mathbf{b}$ es una $\hat{\mathbf{x}}$ en $\mathbb{R}^n$ tal que

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

para toda $\mathbf{x}$ en $\mathbb{R}^n$.

El aspecto más importante del problema de mínimos cuadrados es que no importa cuál $\mathbf{x}$ se elija, el vector $A\mathbf{x}$ necesariamente estará en el espacio de columnas. Así que se busca un $\mathbf{x}$ adecuado para convertir a $A\mathbf{x}$ en el punto de Col A más cercano a $\mathbf{b}$. Vea la figura 1. (Por supuesto, si sucede que $\mathbf{b}$ está en Col A, entonces $\mathbf{b}$ *es* $A\mathbf{x}$ para algún $\mathbf{x}$, y tal $\mathbf{x}$ es una “solución por mínimos cuadrados”.)

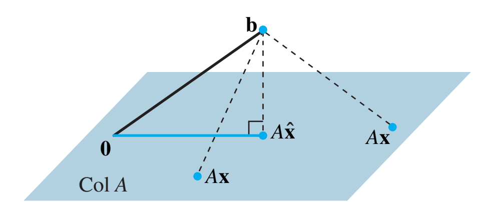


FIGURA 1 El vector $\mathbf{b}$ está más cerca de $A\hat{\mathbf{x}}$ que de $A\mathbf{x}$ para otro $\mathbf{x}$.

Solución del problema general de mínimos cuadrados

Dados los A y $\mathbf{b}$ anteriores, aplique el teorema de aproximación óptima dado en la sección 6.3 al subespacio Col A. Sea

$$\hat{\mathbf{b}} = \text{proy}_{\text{Col}\,A}\,\mathbf{b}$$

Puesto que $\hat{\mathbf{b}}$ está en el espacio de columnas de A, la ecuación $A\mathbf{x} = \hat{\mathbf{b}}$ *es* consistente, y existe un $\hat{\mathbf{x}}$ en $\mathbb{R}^n$ tal que

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}} \tag{1}$$

Puesto que $\hat{\mathbf{b}}$ es el punto de Col A más cercano a $\mathbf{b}$, un vector $\hat{\mathbf{x}}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ si, y sólo si, $\hat{\mathbf{x}}$ satisface (1). Un $\hat{\mathbf{x}}$ tal en $\mathbb{R}^n$ es una lista de pesos que estructurará $\hat{\mathbf{b}}$ a partir de las columnas de A. Vea la figura 2. [Existen muchas soluciones de (1) si la ecuación tiene variables libres.]

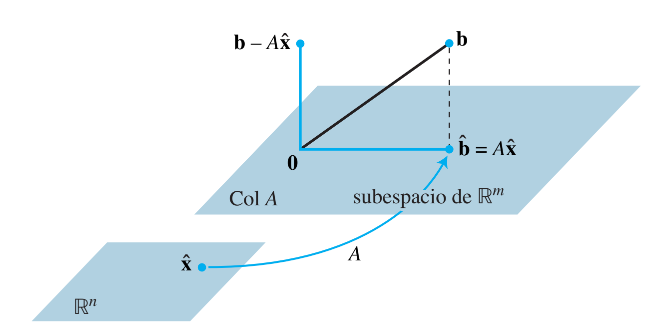


FIGURA 2 La solución por mínimos cuadrados $\hat{\mathbf{x}}$ está en $\mathbb{R}^n$.

Suponga que $\hat{\mathbf{x}}$ satisface $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. De acuerdo con el teorema de la descomposición ortogonal dado en la sección 6.3, la proyección $\mathbf{b}$ tiene la propiedad de que $\mathbf{b} - \hat{\mathbf{b}}$ es ortogonal a Col A, entonces $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a cualquier columna de A. Si $\mathbf{a}_j$

es cualquier columna de A, entonces $\mathbf{a}_j \cdot (\mathbf{b} - A\hat{\mathbf{x}}) = 0$, y $\mathbf{a}_j^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0$. Puesto que cualquier $\mathbf{a}_j^T$ es una fila de A^T,

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0} \tag{2}$$

(Esta ecuación también se deriva del teorema 3 dado en la sección 6.1.) Entonces

$$\begin{aligned} A^T\mathbf{b} - A^TA\hat{\mathbf{x}} &= \mathbf{0} \\ A^TA\hat{\mathbf{x}} &= A^T\mathbf{b} \end{aligned}$$

Estos cálculos muestran que cada solución por mínimos cuadrados de A$\mathbf{x} = \mathbf{b}$ satisface la ecuación

$$A^TA\mathbf{x} = A^T\mathbf{b} \tag{3}$$

La ecuación matricial (3) representa un sistema de ecuaciones lineales llamadas **ecuaciones normales para** $A\mathbf{x} = \mathbf{b}$. Una solución de (3) se denota frecuentemente como $\hat{\mathbf{x}}$.

TEOREMA 13

El conjunto de soluciones por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ coincide con el conjunto no vacío de soluciones de las ecuaciones normales $A^TA\mathbf{x} = A^T\mathbf{b}$.

DEMOSTRACIÓN Como se mostró anteriormente, el conjunto de las soluciones por mínimos cuadrados no es vacío y cualquier $\hat{\mathbf{x}}$ tal cumple con las ecuaciones normales. De manera recíproca, suponga que $\hat{\mathbf{x}}$ satisface $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$. Entonces $\hat{\mathbf{x}}$ satisface a la ecuación (2) anterior, lo cual demuestra que $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a las filas de A^T y, por lo tanto, es ortogonal a las columnas de A. Como las columnas de A generan Col A, el vector $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a toda Col A. Por lo tanto, la ecuación

$$\mathbf{b} = A\hat{\mathbf{x}} + (\mathbf{b} - A\hat{\mathbf{x}})$$

es una descomposición de $\mathbf{b}$ en la suma de un vector de Col A y un vector ortogonal a Col A. De acuerdo con la unicidad de la descomposición ortogonal, $A\hat{\mathbf{x}}$ debe ser la proyección ortogonal de $\mathbf{b}$ sobre Col A. Esto es, $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, y $\hat{\mathbf{x}}$ es una solución por mínimos cuadrados. ■

EJEMPLO 1 Encuentre una solución por mínimos cuadrados del sistema inconsistente $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solución Para usar (3), calcule:

$$A^TA = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Entonces la ecuación $A^TA\mathbf{x} = A^T\mathbf{b}$ se vuelve

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Pueden usarse operaciones por fila para resolver este sistema, pero como A^TA es invertible y 2×2, probablemente sea más fácil calcular

$$(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

y luego resolver $A^TA\mathbf{x} = A^T\mathbf{b}$ como

$$\begin{aligned} \hat{\mathbf{x}} &= (A^TA)^{-1}A^T\mathbf{b} \\ &= \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$

En muchos cálculos, A^TA es invertible, pero no siempre es el caso. El siguiente ejemplo incluye una matriz similar a la que aparece en aquellos problemas de estadística que suelen llamarse problemas *de análisis de varianza*.

EJEMPLO 2 Encuentre una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

Solución Calcule

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

La matriz aumentada para $A^TA\mathbf{x} = A^T\mathbf{b}$ es

$$\begin{bmatrix} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general es $x_1 = 3 - x_4$, $x_2 = -5 + x_4$, $x_3 = -2 + x_4$, y x_4 es libre. Así que la solución general por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ tiene la forma

$$\hat{\mathbf{x}} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

El teorema siguiente proporciona un criterio útil para cuando existe solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. (Por supuesto, la proyección ortogonal $\hat{\mathbf{b}}$ siempre es única.)

TEOREMA 14

La matriz A^TA es invertible si, y sólo si, las columnas de A son linealmente independientes. En este caso, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene solamente una solución por mínimos cuadrados $\hat{\mathbf{x}}$, y está dada por

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}. \qquad (4)$$

En los ejercicios 19 a 21 se incluye un compendio de los elementos principales de una demostración del teorema 14, en el cual también se repasan conceptos del capítulo 4. La fórmula (4) para $\hat{\mathbf{x}}$ es útil especialmente para propósitos teóricos y cálculos a mano cuando A^TA es una matriz invertible de 2×2.

Cuando se utiliza una solución por mínimos cuadrados $\hat{\mathbf{x}}$ para producir $A\hat{\mathbf{x}}$ como una aproximación a $\mathbf{b}$, la distancia de $\mathbf{b}$ a $A\hat{\mathbf{x}}$ se denomina **error de mínimos cuadrados** de esta aproximación.

EJEMPLO 3 Dados A y $\mathbf{b}$ como en el ejemplo 1, determine el error de mínimos cuadrados en la solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

Solución Del ejemplo 1,

$$\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} \quad \text{y} \quad A\hat{\mathbf{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

De aquí que

$$\mathbf{b} - A\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$$

y

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$$

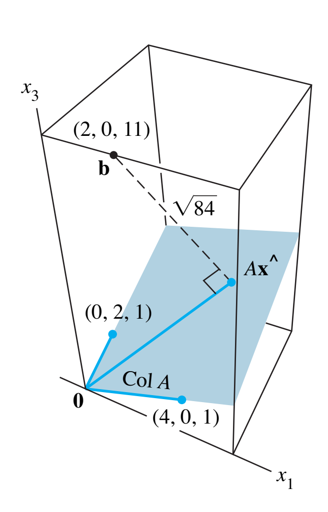


FIGURA 3

El error de mínimos cuadrados es $\sqrt{84}$. Para cualquier $\mathbf{x}$ en $\mathbb{R}^2$, la distancia entre $\mathbf{b}$ y el vector $A\mathbf{x}$ es de al menos $\sqrt{84}$. Véase la figura 3. Observe que la solución por mínimos cuadrados $\hat{\mathbf{x}}$ no aparece en la figura.

Cálculo alternativo de una solución por mínimos cuadrados

En el siguiente ejemplo se muestra cómo encontrar una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando las columnas de A son ortogonales. Tales matrices aparecen a menudo en problemas de regresión lineal, los cuales se verán en la siguiente sección.

EJEMPLO 4 Encuentre una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Solución Como las columnas $\mathbf{a}_1$ y $\mathbf{a}_2$ de A son ortogonales, la proyección ortogonal de $\mathbf{b}$ sobre Col A está dada por

$$\hat{\mathbf{b}} = \frac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1}\mathbf{a}_1 + \frac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2}\mathbf{a}_2 = \frac{8}{4}\mathbf{a}_1 + \frac{45}{90}\mathbf{a}_2 \tag{5}$$

$$= \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ -1 \\ 1/2 \\ 7/2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 5/2 \\ 11/2 \end{bmatrix}$$

Ahora que se conoce $\hat{\mathbf{b}}$, puede resolverse $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Pero esto es trivial, pues ya se sabe qué pesos colocar en las columnas de A para producir $\hat{\mathbf{b}}$. A partir de (5) resulta claro que

$$\hat{\mathbf{x}} = \begin{bmatrix} 8/4 \\ 45/90 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}$$

En algunos casos, la ecuación normal para un problema de mínimos cuadrados puede estar *mal condicionada*; esto es, errores pequeños en los cálculos de las entradas de A^TA pueden ocasionar, a veces, errores relativamente grandes en la solución $\hat{\mathbf{x}}$. Si las columnas de A son linealmente independientes, a menudo puede hacerse un cálculo más confiable de la solución por mínimos cuadrados usando una factorización QR de A (descrita en la sección 6.4).[1]

TEOREMA 15 Dada una matriz A de $m \times n$ con columnas linealmente independientes, sea $A = QR$ una factorización QR de A como en el teorema 12. Entonces, para cada $\mathbf{b}$ en $\mathbb{R}^m$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución única por mínimos cuadrados, dada por

$$\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b} \tag{6}$$

[1]El método QR se compara con el método estándar de ecuaciones normales en G. Golub y C. Van Loan, *Matrix Computations*, 3a. ed. (Baltimore: Johns Hopkins Press, 1996), págs. 230-231.

DEMOSTRACIÓN Sea $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$. Entonces

$$A\hat{\mathbf{x}} = QR\hat{\mathbf{x}} = QRR^{-1}Q^T\mathbf{b} = QQ^T\mathbf{b}$$

De acuerdo con el teorema 12, las columnas de Q forman una base ortonormal para Col A. Así que, según el teorema 10, $QQ^T\mathbf{b}$ es la proyección ortogonal $\hat{\mathbf{b}}$ de $\mathbf{b}$ sobre Col A. Entonces $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, lo cual demuestra que $\hat{\mathbf{x}}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. La unicidad de $\hat{\mathbf{x}}$ se deduce del teorema 14. ■

NOTA NUMÉRICA

Puesto que en el teorema 15 R es triangular superior, $\hat{\mathbf{x}}$ debería calcularse como la solución exacta de la ecuación

$$R\mathbf{x} = Q^T\mathbf{b} \tag{7}$$

Es mucho más rápido resolver (7) por sustitución regresiva u operaciones por fila que calcular R^{-1} y usar (6).

EJEMPLO 5 Encuentre la solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

Solución Puede obtenerse la factorización QR de A como en la sección 6.4.

$$A = QR = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}$$

Entonces

$$Q^T\mathbf{b} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

La solución por mínimos cuadrados $\hat{\mathbf{x}}$ satisface $R\mathbf{x} = Q^T\mathbf{b}$; esto es,

$$\begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

Esta ecuación se resuelve fácilmente y produce $\hat{\mathbf{x}} = \begin{bmatrix} 10 \\ -6 \\ 2 \end{bmatrix}$. ■

PROBLEMAS DE PRÁCTICA

1. Sean $A = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$. Encuentre una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, y calcule el error de mínimos cuadrados asociado.
2. ¿Qué puede decirse acerca de la solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando $\mathbf{b}$ es ortogonal a las columnas de A?

6.5 EJERCICIOS

En los ejercicios 1 a 4, encuentre una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, para ello (a) estructure las ecuaciones normales para $\hat{\mathbf{x}}$, y (b) despeje $\hat{\mathbf{x}}$.

1. $A = \begin{bmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$
2. $A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix}$
3. $A = \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 1 \\ -4 \\ 2 \end{bmatrix}$
4. $A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$

En los ejercicios 5 y 6, describa por mínimos cuadrados todas las soluciones de la ecuación $A\mathbf{x} = \mathbf{b}$.

5. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 8 \\ 2 \end{bmatrix}$
6. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ 2 \\ 3 \\ 6 \\ 5 \\ 4 \end{bmatrix}$
7. Calcule el error de mínimos cuadrados asociado a la solución por mínimos cuadrados obtenida en el ejercicio 3.
8. Calcule el error de mínimos cuadrados asociado a la solución por mínimos cuadrados obtenida en el ejercicio 4.

En los ejercicios 9 a 12, encuentre (a) la proyección ortogonal de $\mathbf{b}$ sobre Col A, y (b) una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

9. $A = \begin{bmatrix} 1 & 5 \\ 3 & 1 \\ -2 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 4 \\ -2 \\ -3 \end{bmatrix}$
10. $A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \\ 1 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$
11. $A = \begin{bmatrix} 4 & 0 & 1 \\ 1 & -5 & 1 \\ 6 & 1 & 0 \\ 1 & -1 & -5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
12. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 6 \end{bmatrix}$
13. Sean $A = \begin{bmatrix} 3 & 4 \\ -2 & 1 \\ 3 & 4 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 11 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$, y $\mathbf{v} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$. Calcule $A\mathbf{u}$ y $A\mathbf{v}$, y compárelos con $\mathbf{b}$. ¿Podría $\mathbf{u}$ ser una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$? (Responda sin calcular una solución por mínimos cuadrados.)
14. Sean $A = \begin{bmatrix} 2 & 1 \\ -3 & -4 \\ 3 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ 4 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$, y $\mathbf{v} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$. Calcule $A\mathbf{u}$ y $A\mathbf{v}$, y compárelos con $\mathbf{b}$. ¿Es posible que al menos $\mathbf{u}$ o $\mathbf{v}$ sean una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$? (Responda sin calcular una solución por mínimos cuadrados.)

En los ejercicios 15 y 16, utilice la factorización $A = QR$ para encontrar la solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

15. $A = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \\ 1/3 & -2/3 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$

16. $A = \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix}$

En los ejercicios 17 y 18, A es una matriz $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

17. a. El problema general de mínimos cuadrados consiste en encontrar una $\mathbf{x}$ que haga $A\mathbf{x}$ tan cercana como sea posible a $\mathbf{b}$.

b. Una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es un vector $\hat{\mathbf{x}}$ que satisface $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, donde $\hat{\mathbf{b}}$ es la proyección ortogonal de $\mathbf{b}$ sobre Col A.

c. Una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es un vector $\hat{\mathbf{x}}$ tal que $\|\mathbf{b} - A\mathbf{x}\| \leq \|\mathbf{b} - A\hat{\mathbf{x}}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

d. Cualquier solución de $A^TA\mathbf{x} = A^T\mathbf{b}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

e. Si las columnas de A son linealmente independientes, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ tiene exactamente una solución por mínimos cuadrados.

18. a. Si $\mathbf{b}$ está en el espacio columna de A, entonces toda solución de $A\mathbf{x} = \mathbf{b}$ es una solución por mínimos cuadrados.

b. La solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es el punto del espacio de columnas de A más cercano a $\mathbf{b}$.

c. Una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es una lista de pesos que, cuando se aplican a las columnas de A, producen la proyección ortogonal de $\mathbf{b}$ sobre Col A.

d. Si $\hat{\mathbf{x}}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, entonces $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

e. Las ecuaciones normales proporcionan siempre un método confiable para calcular soluciones por mínimos cuadrados.

f. Si A tiene una factorización QR, por ejemplo $A = QR$, entonces la mejor manera de encontrar la solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es calcular $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$.

19. Sea A una matriz de $m \times n$. Siga los siguientes pasos para demostrar que un vector $\mathbf{x}$ en $\mathbb{R}^n$ satisface $A\mathbf{x} = \mathbf{0}$ si, y sólo si, $A^TA\mathbf{x} = \mathbf{0}$. Esto demostrará que Nul A = Nul A^TA.

a. Demuestre que si $A\mathbf{x} = \mathbf{0}$, entonces $A^TA\mathbf{x} = \mathbf{0}$.

b. Suponga que $A^TA\mathbf{x} = \mathbf{0}$. Explique por qué $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{0}$, y utilice esto para mostrar que $A\mathbf{x} = \mathbf{0}$.

20. Sea A una matriz de $m \times n$ tal que A^TA es invertible. Demuestre que las columnas de A son linealmente independientes. [*Cuidado:* No puede suponerse que A sea invertible; podría no ser siquiera cuadrada.]

21. Sea A una matriz de $m \times n$ cuyas columnas son linealmente independientes. [*Cuidado: A* no necesita ser cuadrada.]

a. Utilice el ejercicio 19 para mostrar que A^TA es una matriz invertible.

b. Explique por qué A debe tener por lo menos tantas filas como columnas.

c. Determine el rango de A.

22. Utilice el ejercicio 19 para mostrar que rango A^TA = rango A. [*Pista:* ¿Cuántas columnas tiene A^TA? ¿Cómo se relaciona esto con el rango de A^TA?]

23. Suponga que A es de $m \times n$ con columnas linealmente independientes y que $\mathbf{b}$ está en $\mathbb{R}^m$. Utilice las ecuaciones normales y produzca una fórmula para $\hat{\mathbf{b}}$, la proyección de $\mathbf{b}$ sobre Col A. [*Sugerencia:* Primero encuentre $\hat{\mathbf{x}}$. La fórmula no requiere una base ortogonal para Col A.]

24. Encuentre una fórmula para la solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando las columnas de A son ortonormales.

25. Describa todas las soluciones por mínimos cuadrados del sistema

$$\begin{aligned} x + y &= 2 \\ x + y &= 4 \end{aligned}$$

26. **[M]** El ejemplo 3 de la sección 4.8 mostró un filtro lineal pasabajas que convierte una señal $\{y_k\}$ en $\{y_{k+1}\}$ y transforma una señal de frecuencia más alta $\{w_k\}$ en la señal cero, donde $y_k = \cos(\pi k/4)$ y $w_k = \cos(3\pi k/4)$. Los cálculos siguientes diseñan un filtro con aproximadamente esas propiedades. La ecuación del filtro es

$$a_0 y_{k+2} + a_1 y_{k+1} + a_2 y_k = z_k \qquad \text{para toda } k \tag{8}$$

Como las señales son periódicas, con periodo 8, es suficiente con estudiar la ecuación (8) para $k = 0, \ldots, 7$. La acción ejercida sobre las dos señales descritas anteriormente se traduce en dos conjuntos de ocho ecuaciones, los cuales se muestran a continuación y en la página 418:

$$\begin{array}{l} k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array} \overset{\begin{array}{ccc} y_{k+2} & y_{k+1} & y_k \end{array}}{\begin{bmatrix} 0 & .7 & 1 \\ -.7 & 0 & .7 \\ -1 & -.7 & 0 \\ -.7 & -1 & -.7 \\ 0 & -.7 & -1 \\ .7 & 0 & -.7 \\ 1 & .7 & 0 \\ .7 & 1 & .7 \end{bmatrix}} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \overset{y_{k+1}}{\begin{bmatrix} .7 \\ 0 \\ -.7 \\ -1 \\ -.7 \\ 0 \\ .7 \\ 1 \end{bmatrix}};$$

$$\begin{array}{l} k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array} \overset{\begin{array}{ccc} w_{k+2} & w_{k+1} & w_k \end{array}}{\begin{bmatrix} 0 & -.7 & 1 \\ .7 & 0 & -.7 \\ -1 & .7 & 0 \\ .7 & -1 & .7 \\ 0 & .7 & -1 \\ -.7 & 0 & .7 \\ 1 & -.7 & 0 \\ -.7 & 1 & -.7 \end{bmatrix}} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Escriba una ecuación $A\mathbf{x} = \mathbf{b}$, donde A es una matriz de 16 $\times$ 3 formada con las dos anteriores matrices de coeficientes, y donde $\mathbf{b}$ en $\mathbb{R}^{16}$ está formado a partir de los dos miembros derechos de las ecuaciones. Encuentre las a_0, a_1, a_2 dadas mediante la solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. (El .7 de los datos anteriores se usó como una aproximación de $\sqrt{2}/2$, para ilustrar cómo se podría llevar a cabo un cálculo típico en un problema aplicado. Si en su lugar se usara .707, los coeficientes del filtro resultantes coincidirían en por lo menos siete posiciones decimales con $\sqrt{2}/4$, $1/2$, y $\sqrt{2}/4$, como los valores producidos mediante cálculos aritméticos exactos.)

CD Mínimos cuadrados y QR (Least-Squares and QR)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Primero, calcule

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 0 \\ 9 & 83 & 28 \\ 0 & 28 & 14 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} -3 \\ -65 \\ -28 \end{bmatrix}$$

Enseguida, reduzca por filas la matriz aumentada para las ecuaciones normales: $A^TA\mathbf{x} = A^T\mathbf{b}$:

$$\begin{bmatrix} 3 & 9 & 0 & -3 \\ 9 & 83 & 28 & -65 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 56 & 28 & -56 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -3/2 & 2 \\ 0 & 1 & 1/2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general por mínimos cuadrados es $x_1 = 2 + \frac{3}{2}x_3$, $x_2 = -1 - \frac{1}{2}x_3$, con x_3 libre. Para una solución específica, tome $x_3 = 0$ (por ejemplo), y obtenga

$$\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

Para encontrar el error de mínimos cuadrados, calcule

$$\hat{\mathbf{b}} = A\hat{\mathbf{x}} = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$$

Resulta que $\hat{\mathbf{b}} = \mathbf{b}$, entonces $\|\mathbf{b} - \hat{\mathbf{b}}\| = 0$. El error de mínimos cuadrados es cero porque sucede que $\mathbf{b}$ está en Col A.

2. Si $\mathbf{b}$ es ortogonal a las columnas de A, entonces la proyección de $\mathbf{b}$ sobre el espacio columna de A es $\mathbf{0}$. En este caso, una solución por mínimos cuadrados $\hat{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ satisface $A\hat{\mathbf{x}} = \mathbf{0}$.

6.6 APLICACIONES A MODELOS LINEALES

Una tarea común en ciencias e ingeniería es analizar y comprender las relaciones presentes entre diversas cantidades que varían. En esta sección se describirán diversas situaciones en las que se usan datos para estructurar o verificar una fórmula que predice el valor de una variable como una función de otras variables. En cada caso, el problema equivaldrá a resolver un problema de mínimos cuadrados.

Para facilitar la aplicación del análisis a problemas reales que los lectores podrán encontrar posteriormente en el desarrollo profesional de su carrera, se elige la notación que suele utilizarse en el análisis estadístico de datos científicos y de ingeniería. En lugar de $A\mathbf{x} = \mathbf{b}$, se escribe $X\boldsymbol{\beta} = \mathbf{y}$ y se llama a X **matriz de diseño**, $\boldsymbol{\beta}$ es el **vector parámetro** y $\mathbf{y}$ el **vector de observación**.

Líneas de mínimos cuadrados

La relación más sencilla entre dos variables x y y es la ecuación lineal $y = \beta_0 + \beta_1 x$.[1] Los datos experimentales a menudo producen puntos $(x_1, y_1), \ldots, (x_n, y_n)$ que, al graficarse, parecen quedar cerca de una línea. Se desea determinar los parámetros β_0 y β_1 que dejen a la línea tan "cercana" a los puntos como sea posible.

Suponga que β_0 y β_1 están fijos, y considere la línea $y = \beta_0 + \beta_1 x$ que aparece en la figura 1. Para cada punto de los datos (x_j, y_j) hay un punto correspondiente $(x_j, \beta_0 + \beta_1 x_j)$ sobre la línea con la misma coordenada x. A y_j se le llama el *valor observado* de y, y a $\beta_0 + \beta_1 x_j$ el valor *pronosticado* de y (determinado mediante la línea). La diferencia entre un valor de y observado y uno pronosticado se denomina *residual.*

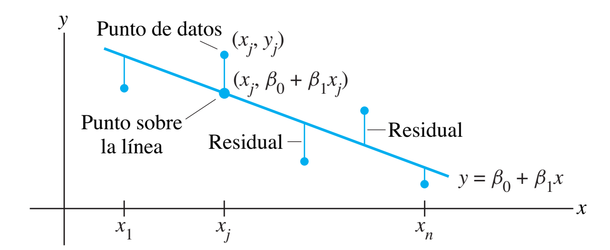


FIGURA 1 Ajuste de una línea a datos experimentales.

Existen varias maneras de medir qué tan "cercana" está la línea a los datos. La elección acostumbrada (que se elige principalmente por la sencillez de los cálculos matemáticos) es sumar los cuadrados de los residuales. La **línea de mínimos cuadrados** es la línea $y = \beta_0 + \beta_1 x$ que minimiza la suma de los cuadrados de los residuales. Esta línea también se conoce como **línea de regresión de y sobre x**, porque se supone que cualquier error en los datos está únicamente en las coordenadas de y. Los coeficientes β_0, β_1 de la línea se llaman **coeficientes de regresión** (lineal).[2]

---

[1]Esta notación se usa comúnmente para líneas de mínimos cuadrados en lugar de $y = mx + b$.

[2]Si los errores de medición están en x en vez de en y, simplemente se intercambian las coordenadas de los datos (x_j, y_j) antes de graficar los puntos y calcular la línea de regresión. Si ambas coordenadas están sujetas a posibles errores, entonces se puede elegir la línea que minimice la suma de los cuadrados de las distancias *ortogonales* (perpendiculares) desde los puntos hasta la línea. Vea los problemas de práctica de la sección 7.5.

Si los puntos de datos estuvieran sobre la línea, los parámetros β_0 y β_1 satisfarían las ecuaciones

Valor de y pronosticado		Valor de y observado
$\beta_0 + \beta_1 x_1$	$=$	y_1
$\beta_0 + \beta_1 x_2$	$=$	y_2
$\vdots$		$\vdots$
$\beta_0 + \beta_1 x_n$	$=$	y_n

Este sistema puede escribirse como

$$X\boldsymbol{\beta} = \mathbf{y}, \quad \text{donde } X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{1}$$

Por supuesto, si los puntos de datos no están sobre una línea, entonces no hay parámetros β_0, β_1 para los cuales los valores de y pronosticados en $X\boldsymbol{\beta}$ sean iguales a los valores de y observados en $\mathbf{y}$, y $X\boldsymbol{\beta} = \mathbf{y}$ no tiene solución. ¡Éste es un problema de mínimos cuadrados, $A\mathbf{x} = \mathbf{b}$, con una notación diferente!

El cuadrado de la distancia entre los vectores $X\boldsymbol{\beta}$ y $\mathbf{y}$ es precisamente la suma de los cuadrados de los residuales. El $\boldsymbol{\beta}$ que minimiza esta suma también minimiza la distancia entre $X\boldsymbol{\beta}$ y $\mathbf{y}$. *El cálculo de la solución por mínimos cuadrados de* $X\boldsymbol{\beta} = \mathbf{y}$ *equivale a encontrar el* $\boldsymbol{\beta}$ *que determina la línea de mínimos cuadrados de la figura 1.*

EJEMPLO 1 Encuentre la ecuación $y = \beta_0 + \beta_1 x$ de la línea de mínimos cuadrados que se ajuste mejor a los puntos de datos (2, 1), (5, 2), (7, 3), (8, 3).

Solución Utilice las coordenadas x de los datos para estructurar la matriz X en (1) y las coordenadas y para estructurar el vector $\mathbf{y}$:

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

Para la solución por mínimos cuadrados de $X\boldsymbol{\beta} = \mathbf{y}$, obtenga las ecuaciones normales (con la notación nueva):

$$X^T X\boldsymbol{\beta} = X^T\mathbf{y}$$

Es decir, calcule

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Las ecuaciones normales son

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

De donde

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

Entonces la línea de mínimos cuadrados tiene la ecuación

$$y = \frac{2}{7} + \frac{5}{14}x$$

Vea la figura 2.

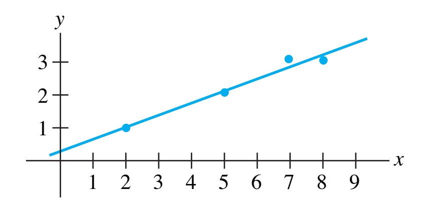


FIGURA 2 La línea de mínimos cuadrados $y = \frac{2}{7} + \frac{5}{14}x$.

Una práctica común antes de calcular una línea de mínimos cuadrados es calcular el promedio $\overline{x}$ de los valores x originales y formar una nueva variable $x^* = x - \overline{x}$. Se dice que los nuevos datos x están en **forma de desviación media**. En este caso, las dos columnas de la matriz de diseño serán ortogonales. La solución de las ecuaciones normales se simplifica, como en el ejemplo 4 de la sección 6.5. Vea los ejercicios 17 y 18.

El modelo lineal general

En algunas aplicaciones, es necesario ajustar los puntos de datos a algo distinto a una línea recta. En los ejemplos siguientes, la ecuación matricial aún es $X\boldsymbol{\beta} = \mathbf{y}$, pero la forma específica de X cambia de un problema al siguiente. Los estadísticos introducen generalmente un **vector residual** $\boldsymbol{\epsilon}$, definido mediante $\boldsymbol{\epsilon} = \mathbf{y} - X\boldsymbol{\beta}$, y escriben

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Cualquier ecuación de esta forma se denomina **modelo lineal**. Una vez que X y $\mathbf{y}$ se determinan, el objetivo es minimizar la longitud de $\boldsymbol{\epsilon}$, lo cual equivale a encontrar una

solución por mínimos cuadrados de $X\boldsymbol{\beta} = \mathbf{y}$. En todo caso, la solución por mínimos cuadrados $\hat{\boldsymbol{\beta}}$ es una solución de las ecuaciones normales

$$X^T X \boldsymbol{\beta} = X^T \mathbf{y}$$

Ajuste de otras curvas por mínimos cuadrados

Cuando los puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$ de un "diagrama de dispersión" no están cerca de ninguna línea, podría ser adecuado postular alguna otra relación funcional entre x y y.

Los tres ejemplos siguientes muestran cómo ajustar los datos por medio de curvas que tienen la forma general

$$y = \beta_0 f_0(x) + \beta_1 f_1(x) + \cdots + \beta_k f_k(x) \qquad (2)$$

donde las $f_0, \ldots, f_k$ son funciones conocidas y los $\beta_0, \ldots, \beta_k$ son parámetros que deben determinarse. Como se verá, la ecuación (2) describe un modelo lineal porque es lineal en los parámetros desconocidos.

Para un valor específico de x, (2) proporciona un valor pronosticado o "ajustado" de y. La diferencia entre el valor observado y el valor pronosticado es el residual. Los parámetros $\beta_0, \ldots, \beta_k$ deben determinarse de manera que minimicen la suma de los cuadrados de los residuales.

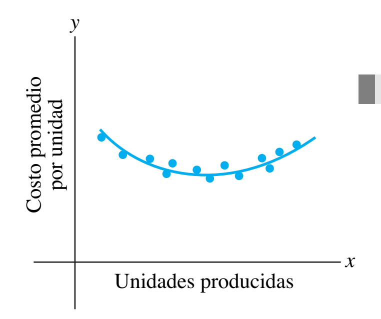


FIGURA 3 Curva de costo promedio.

EJEMPLO 2 Suponga que los puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$ parecen situarse sobre algún tipo de parábola en lugar de en línea recta. Por ejemplo, si la coordenada x denota el nivel de producción de una compañía y y denota el costo promedio por unidad de operación en un nivel de x unidades por día, entonces una curva típica de costos promedio semeja una parábola que se abre hacia arriba (figura 3). En ecología se usa una curva parabólica abierta hacia abajo para modelar la producción primaria neta de nutrimentos en una planta en función del área superficial del follaje (figura 4). Suponga que se desea aproximar los datos mediante una ecuación de la forma

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 \qquad (3)$$

Describa el modelo lineal que produce un "ajuste por mínimos cuadrados" de los datos mediante la ecuación (3).

Solución La ecuación (3) describe la relación ideal. Suponga que los valores reales de los parámetros son $\beta_0, \beta_1, \beta_2$. Entonces las coordenadas del primer punto de datos (x_1, y_1) satisfacen una ecuación de la forma

$$y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \epsilon_1$$

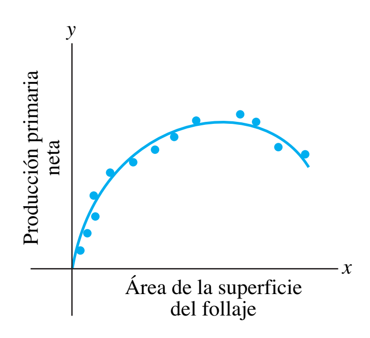


FIGURA 4 Producción de nutrimentos.

donde ϵ_1 es el error residual entre el valor observado y_1 y el valor y pronosticado $\beta_0 + \beta_1 x_1 + \beta_2 x_1^2$. Cada uno de los puntos de datos determina una ecuación similar:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \epsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \epsilon_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_n + \beta_2 x_n^2 + \epsilon_n \end{aligned}$$

Es una cuestión sencilla escribir este sistema de ecuaciones en la forma $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$. Para encontrar X, se inspeccionan algunas de las primeras filas del sistema y se busca el patrón.

$$\underset{\mathbf{y}}{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}} = \underset{X}{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}} \underset{\boldsymbol{\beta}}{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}} + \underset{\boldsymbol{\epsilon}}{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}}$$

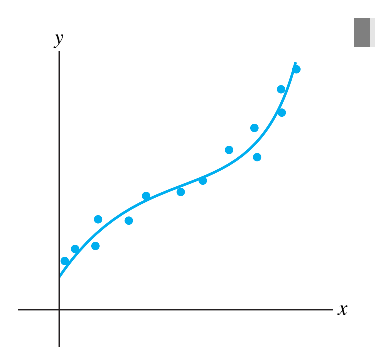


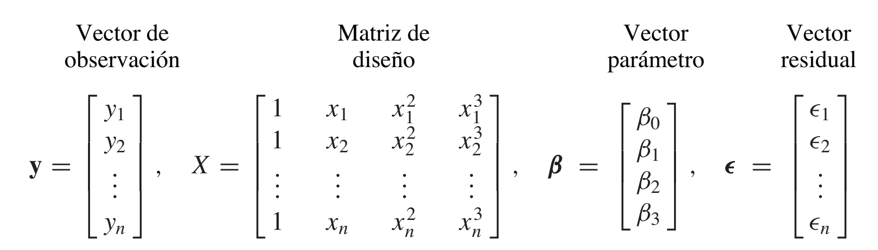
FIGURA 5 Puntos de datos sobre una curva cúbica.

EJEMPLO 3 Si los puntos de datos tienden a seguir un patrón como el de la figura 5, entonces un modelo apropiado podría tener una ecuación de la forma

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

Datos así podrían provenir, por ejemplo, de los costos totales de una compañía, como una función del nivel de producción. Describa el modelo lineal que proporciona un ajuste por mínimos cuadrados de este tipo a datos como $(x_1, y_1), \dots, (x_n, y_n)$.

Solución Por medio de un análisis semejante al del ejemplo 2, se obtiene



$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Regresión múltiple

Suponga que cierto experimento implica dos variables independientes —por ejemplo, u y v— y una variable dependiente, y. Una ecuación sencilla para predecir y a partir de u y v tiene la forma

$$y = \beta_0 + \beta_1 u + \beta_2 v \tag{4}$$

Una ecuación más general de predicción podría tener la forma

$$y = \beta_0 + \beta_1 u + \beta_2 v + \beta_3 u^2 + \beta_4 uv + \beta_5 v^2 \tag{5}$$

Esta ecuación se usa en geología, por ejemplo, para modelar superficies de erosión, circos glaciales, pH del suelo y otras cantidades. En tales casos, el ajuste por mínimos cuadrados se denomina *superficie de tendencia*.

Tanto (4) como (5) llevan a un modelo lineal porque son lineales en los parámetros desconocidos (aunque u y v estén multiplicadas). En general, surgirá un modelo lineal siempre que se deba predecir y a partir de una ecuación de la forma

$$y = \beta_0 f_0(u, v) + \beta_1 f_1(u, v) + \cdots + \beta_k f_k(u, v)$$

donde $f_0, \dots, f_k$ son cualquier tipo de funciones conocidas y $\beta_0, \dots, \beta_k$ son pesos desconocidos.

EJEMPLO 4 En geografía, se estructuran modelos locales de terreno a partir de datos $(u_1, v_1, y_1), \ldots, (u_n, v_n, y_n)$, donde u_j, v_j y y_j son latitud, longitud y altitud, respectivamente. Describa el modelo lineal basado en (4) que proporciona un ajuste por mínimos cuadrados a datos de este tipo. La solución es el *plano de mínimos cuadrados*. Vea la figura 6.

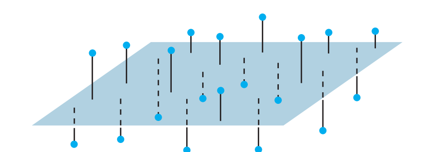


FIGURA 6 Un plano de mínimos cuadrados.

Solución Se espera que los datos satisfagan las siguientes ecuaciones:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 u_1 + \beta_2 v_1 + \epsilon_1 \\ y_2 &= \beta_0 + \beta_1 u_2 + \beta_2 v_2 + \epsilon_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 u_n + \beta_2 v_n + \epsilon_n \end{aligned}$$

Este sistema tiene la forma matricial $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde

Vector observación	Matriz de diseño	Vector parámetro	Vector residual

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & u_1 & v_1 \\ 1 & u_2 & v_2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & v_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

El ejemplo 4 muestra que el modelo lineal para la regresión múltiple tiene la misma forma abstracta que el modelo para la regresión simple de los ejemplos anteriores. El álgebra lineal permite comprender el principio general en que se basan todos los modelos lineales. Una vez que X está definida adecuadamente, las ecuaciones normales para $\boldsymbol{\beta}$ tienen la misma forma matricial, sin importar cuántas variables intervengan. Entonces, para cualquier modelo lineal donde X^TX sea invertible, la $\hat{\boldsymbol{\beta}}$ de mínimos cuadrados está dada por $(X^TX)^{-1}X^T\mathbf{y}$.

SG La geometría de un modelo lineal 6 a 19 (The Geometry of a Linear Model 6-19)

Lecturas adicionales

Ferguson, J., *Introduction to Linear Algebra in Geology* (Nueva York: Chapman & Hall, 1994).

Krumbein, W. C. y F. A. Graybill, *An Introduction to Statistical Models in Geology* (Nueva York: McGraw-Hill, 1965).

Legendre, P. y L. Legendre, *Numerical Ecology* (Amsterdam: Elsevier, 1998).

Unwin, David J., *An Introduction to Trend Surface Analysis*, Concepts and Techniques in Modern Geography, núm. 5 (Norwich, Inglaterra: Geo Books, 1975).

PROBLEMA DE PRÁCTICA

Cuando las ventas mensuales de un producto están sujetas a fluctuaciones de temporada, una curva que aproxime los datos de ventas podría tener la forma

$$y = \beta_0 + \beta_1 x + \beta_2 \text{ sen } (2\pi x/12)$$

donde x es el tiempo en meses. El término $\beta_0 + \beta_1 x$ proporciona la tendencia básica de las ventas, y el término seno refleja los cambios por temporada de las ventas. Proporcione la matriz de diseño y el vector de parámetros apropiados para el modelo lineal que conduzca a un ajuste de la ecuación anterior por mínimos cuadrados. Suponga que los datos son $(x_1, y_1), \ldots, (x_n, y_n)$.

6.6 EJERCICIOS

En los ejercicios 1 a 4, encuentre la ecuación $y = \beta_0 + \beta_1 x$ de la línea de mínimos cuadrados que mejor se ajuste a los puntos de datos proporcionados.

1. (0, 1), (1, 1), (2, 2), (3, 2)

2. (1, 0), (2, 1), (4, 2), (5, 3)

3. (−1, 0), (0, 1), (1, 2), (2, 4)

4. (2, 3), (3, 2), (5, 1), (6, 0)

5. Sea X la matriz de diseño usada para encontrar la línea de mínimos cuadrados que se ajuste a los datos $(x_1, y_1), \ldots, (x_n, y_n)$. Utilice un teorema de la sección 6.5 para mostrar que las ecuaciones normales tienen una solución única si, y sólo si, los datos incluyen por lo menos dos puntos de datos con coordenadas x diferentes.

6. Sea X la matriz de diseño del ejemplo 2 correspondiente al ajuste por mínimos cuadrados de una parábola a los datos $(x_1, y_1), \ldots, (x_n, y_n)$. Suponga que x_1, x_2, x_3 son distintos. Explique por qué sólo hay una parábola con un mejor ajuste a los datos, en un sentido de mínimos cuadrados. (Vea el ejercicio 5.)

7. Cierto experimento produce los datos (1, 1.8), (2, 2.7), (3, 3.4), (4, 3.8) y (5, 3.9). Describa el modelo que produce un ajuste por mínimos cuadrados a estos puntos mediante una función de la forma

$$y = \beta_1 x + \beta_2 x^2$$

Dicha función podría surgir, por ejemplo, como las ganancias por la venta de x unidades de un producto, cuando la cantidad ofrecida en venta afecta el precio que se asignará al producto.

a. Proporcione la matriz de diseño, el vector de observación, y el vector de parámetro desconocido.

b. **[M]** Encuentre la curva de mínimos cuadrados asociada para los datos.

8. Una curva sencilla que a menudo es un buen modelo para los costos variables de una compañía, como función del nivel de ventas x, tiene la forma $y = \beta_1 x + \beta_2 x^2 + \beta_3 x^3$. No existe término constante porque no están incluidos los costos fijos.

a. Proporcione la matriz de diseño y el vector de parámetro para el modelo lineal que conduzca a un ajuste por mínimos cuadrados de la ecuación anterior, con los datos $(x_1, y_1), \ldots, (x_n, y_n)$.

b. **[M]** Encuentre la curva de mínimos cuadrados de la forma anterior que se ajuste a los datos (4, 1.58), (6, 2.08), (8, 2.5), (10, 2.8), (12, 3.1), (14, 3.4), (16, 3.8), (18, 4.32), con valores en millares. De ser posible, trace una gráfica que muestre los puntos de datos y la curva de la aproximación cúbica.

9. Cierto experimento produce los datos (1, 7.9), (2, 5.4) y (3, −.9). Describa el modelo que produce un ajuste por mínimos cuadrados a estos puntos mediante una función de la forma

$$y = A \cos x + B \text{ sen } x$$

10. Suponga que las sustancias radiactivas A y B tienen coeficientes de decaimiento de .02 y .07, respectivamente. Si una mezcla de estas dos sustancias en el tiempo $t = 0$ contiene M_A gramos de A y M_B gramos de B, entonces un modelo para la cantidad total y de la mezcla presente en el tiempo t es

$$y = M_A e^{-.02t} + M_B e^{-.07t} \qquad (6)$$

Suponga que las cantidades iniciales M_A, M_B se desconocen, pero que un científico puede medir la cantidad total presente en diferentes tiempos y registra los siguientes puntos (t_i, y_i): (10, 21.34), (11, 20.68), (12, 20.05), (14, 18.87) y (15, 18.30).

a. Describa un modelo lineal que pueda usarse para estimar M_A y M_B.

b. **[M]** Encuentre la curva de mínimos cuadrados basada en (6).

El cometa Halley apareció por última vez en 1986 y reaparecerá en 2061.

11. **[M]** Según la primera ley de Kepler, un cometa debe tener una órbita elíptica, parabólica o hiperbólica (ignorando la atracción gravitacional de los planetas). En coordenadas polares adecuadas, la posición (r, ϑ) de un cometa satisface una ecuación de la forma

$r = \beta - e(r \cdot \cos \vartheta)$

donde β es una constante y e es la *excentricidad* de la órbita, con $0 \leq e < 1$ para una elipse, $e = 1$ para una parábola, y $e > 1$ para una hipérbola. Suponga que las observaciones de un cometa recientemente descubierto proporcionan los datos siguientes. Determine el tipo de órbita y pronostique dónde estará el cometa cuando $\vartheta = 4.6$ (radianes).[2]

ϑ	.88	1.10	1.42	1.77	2.14
r	3.00	2.30	1.65	1.25	1.01

12. **[M]** La presión arterial sistólica de un niño sano (en milímetros de mercurio) y el peso w (en libras) se relacionan aproximadamente por la ecuación

$\beta_0 + \beta_1 \ln w = p$

Utilice los siguientes datos experimentales para estimar la presión arterial sistólica de un niño sano que pesa 100 libras.

w	44	61	81	113	131
$\ln w$	3.78	4.11	4.41	4.73	4.88
p	91	98	103	110	112

13. **[M]** Para medir el desempeño de un avión durante el despegue, se midió su posición horizontal cada segundo, desde $t = 0$ hasta $t = 12$. Las posiciones (en pies) fueron: 0, 8.8, 29.9, 62.0, 104.7, 159.1, 222.0, 294.5, 380.4, 471.1, 571.7, 686.8, 809.2.

a. Encuentre la curva cúbica de mínimos cuadrados $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ para estos datos.

b. Utilice el resultado de (a) para estimar la velocidad del avión cuando $t = 4.5$ segundos.

14. Sean $\overline{x} = \frac{1}{n}(x_1 + \cdots + x_n)$ y $\overline{y} = \frac{1}{n}(y_1 + \cdots + y_n)$. Muestre que la línea de mínimos cuadrados para los datos $(x_1, y_1), \ldots, (x_n, y_n)$ debe pasar por $(\overline{x}, \overline{y})$. Esto es, muestre que $\overline{x}$ y $\overline{y}$ satisfacen la ecuación lineal $\overline{y} = \hat{\beta}_0 + \hat{\beta}_1 \overline{x}$. [*Sugerencia:* Derive esta ecuación de la ecuación vectorial $\mathbf{y} = X\hat{\boldsymbol{\beta}} + \boldsymbol{\epsilon}$. Denote la primera columna de X mediante $\mathbf{1}$. Utilice el hecho de que el vector residual $\boldsymbol{\epsilon}$ es ortogonal al espacio de columnas de X y, por ende, ortogonal a $\mathbf{1}$.]

Dados los datos para un problema de mínimos cuadrados, $(x_1, y_1), \ldots, (x_n, y_n)$, las siguientes abreviaturas resultan muy útiles:

$$\sum x = \sum_{i=1}^n x_i, \quad \sum x^2 = \sum_{i=1}^n x_i^2,$$
$$\sum y = \sum_{i=1}^n y_i, \quad \sum xy = \sum_{i=1}^n x_i y_i$$

Las ecuaciones normales para una línea de mínimos cuadrados $y = \hat{\beta}_0 + \hat{\beta}_1 x$ pueden escribirse en la forma

$$\begin{aligned} n\hat{\beta}_0 + \hat{\beta}_1 \textstyle\sum x &= \textstyle\sum y \\ \hat{\beta}_0 \textstyle\sum x + \hat{\beta}_1 \textstyle\sum x^2 &= \textstyle\sum xy \end{aligned} \tag{7}$$

15. Deduzca las ecuaciones normales (7) a partir de la forma matricial dada en esta sección.

16. Use una matriz inversa para resolver el sistema de ecuaciones (7) y obtener así fórmulas para $\hat{\beta}_0$ y $\hat{\beta}_1$, las cuales aparecen en muchos textos de estadística.

17. a. Vuelva a escribir los datos del ejemplo 1 con nuevas coordenadas x en forma de desviación media. Sea X la matriz de diseño asociada. ¿Por qué son ortogonales las columnas de X?

b. Escriba las ecuaciones normales para los datos del inciso (a), y resuélvalas para encontrar la línea de mínimos cuadrados, $y = \beta_0 + \beta_1 x^*$, donde $x^* = x - 5.5$.

18. Suponga que las coordenadas x de los datos $(x_1, y_1), \ldots, (x_n, y_n)$ están en forma de desviación media, de manera que

[2]La idea básica del ajuste por mínimos cuadrados a los datos se debe a K. F. Gauss (e independientemente, a A. Legendre), cuya fama comenzó a crecer en 1801 cuando utilizó el método para determinar la trayectoria del asteroide *Ceres.* Cuarenta días después de haberse descubierto el asteroide, desapareció tras el Sol. Gauss pronosticó que aparecería diez meses después y dio su posición. Lo preciso del pronóstico sorprendió a la comunidad científica europea.

$\sum x_i = 0$. Muestre que si X es la matriz de diseño para la línea de mínimos cuadrados en este caso, entonces X^TX es una matriz diagonal.

Los ejercicios 19 y 20 involucran una matriz de diseño X con dos o más columnas y una solución por mínimos cuadrados $\hat{\boldsymbol{\beta}}$ de $\mathbf{y} = X\boldsymbol{\beta}$. Considere los siguientes números.

(i) $\|X\hat{\boldsymbol{\beta}}\|^2$ —la suma de los cuadrados del "término de regresión"—. Denote este número mediante SS(R).

(ii) $\|\mathbf{y} - X\hat{\boldsymbol{\beta}}\|^2$ —la suma de los cuadrados del término de error—. Denote este número con SS(E).

(iii) $\|\mathbf{y}\|^2$ —la suma "total" de los cuadrados de los valores y—. Denote este número como SS(T).

Todo texto de estadística que trate acerca de la regresión y el modelo lineal $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ introduce estos números, aunque la terminología y la notación pueden variar un poco. Para simplificar las cosas, suponga que la media de los valores y es cero. En este caso, SS(T) es proporcional a lo que se conoce como la *varianza* del conjunto de valores y.

19. Justifique la ecuación SS(T) = SS(R) + SS(E). [*Sugerencia:* Utilice un teorema y explique por qué se satisfacen las hipótesis del teorema.] Esta ecuación es extremadamente importante en estadística, tanto en la teoría de regresión como en el análisis de varianza.

20. Muestre que $\|X\hat{\boldsymbol{\beta}}\|^2 = \hat{\boldsymbol{\beta}}^TX^T\mathbf{y}$. [*Sugerencia:* Vuelva a escribir el miembro izquierdo y utilice el hecho de que $\hat{\boldsymbol{\beta}}$ satisface las ecuaciones normales.] Esta fórmula para SS(R) se usa en estadística. A partir de esto y del ejercicio 19, obtenga la fórmula estándar para SS(E):

$$\text{SS(E)} = \mathbf{y}^T\mathbf{y} - \hat{\boldsymbol{\beta}}^TX^T\mathbf{y}$$

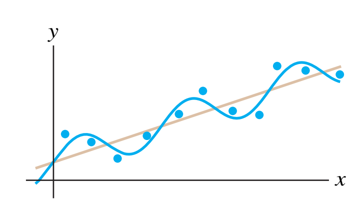
Solución al problema de práctica

X y $\boldsymbol{\beta}$ deben estructurarse de manera que la fila k-ésima de $X\boldsymbol{\beta}$ sea el valor de y pronosticado que corresponde al punto de dato (x_k, y_k), a saber,

$$\beta_0 + \beta_1 x_k + \beta_2 \operatorname{sen}(2\pi x_k/12)$$

Debe quedar claro que

$$X = \begin{bmatrix} 1 & x_1 & \operatorname{sen}(2\pi x_1/12) \\ \vdots & \vdots & \vdots \\ 1 & x_n & \operatorname{sen}(2\pi x_n/12) \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$



Tendencia de las ventas con fluctuaciones de temporada.

6.7 ESPACIOS CON PRODUCTO INTERIOR

Los conceptos de longitud, distancia y ortogonalidad a menudo son importantes en aplicaciones donde interviene un espacio vectorial. Para $\mathbb{R}^n$, estos conceptos se basaban en las propiedades del producto interior listadas en el teorema 1 de la sección 6.1. Para otros espacios, se necesitan productos análogos al producto interior con las mismas propiedades. En la siguiente definición, las conclusiones del teorema 1 se convierten en *axiomas*.

DEFINICIÓN Un **producto interior** dentro de un espacio vectorial V es una función que asocia a cada par de vectores **u** y **v** en V un número real $\langle \mathbf{u}, \mathbf{v} \rangle$, y satisface los siguientes axiomas para todos **u**, **v**, **w** en V y para todo escalar c:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ y $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ si, y sólo si, $\mathbf{u} = \mathbf{0}$

Un espacio vectorial con un producto interior se llama **espacio con producto interior**.

El espacio vectorial $\mathbb{R}^n$ con el producto interior estándar es un espacio con producto interior, y casi todo lo que se explique en este capítulo para $\mathbb{R}^n$ es aplicable a los espacios con producto interior. Los ejemplos de esta sección y de la siguiente establecen la base apropiada para abordar una amplia gama de aplicaciones que se tratan en cursos de ingeniería, física, matemáticas y estadística.

EJEMPLO 1 Fije cualesquiera dos números positivos —por ejemplo, 4 y 5— y, para los vectores $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$ en $\mathbb{R}^2$, sea

$$\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 5u_2v_2 \tag{1}$$

Muestre que (1) define un producto interior.

Solución Desde luego que se satisface el axioma 1, pues $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2 = \langle \mathbf{v}, \mathbf{u} \rangle$. Si $\mathbf{w} = (w_1, w_2)$, entonces

$$\begin{aligned} \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle &= 4(u_1 + v_1)w_1 + 5(u_2 + v_2)w_2 \\ &= 4u_1w_1 + 5u_2w_2 + 4v_1w_1 + 5v_2w_2 \\ &= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \end{aligned}$$

Esto verifica el axioma 2. Para el axioma 3, se tiene que

$$\langle c\mathbf{u}, \mathbf{v} \rangle = 4(cu_1)v_1 + 5(cu_2)v_2 = c(4u_1v_1 + 5u_2v_2) = c\langle \mathbf{u}, \mathbf{v} \rangle$$

Para el axioma 4, observe que $\langle \mathbf{u}, \mathbf{u} \rangle = 4u_1^2 + 5u_2^2 \geq 0$, y que $4u_1^2 + 5u_2^2 = 0$ sólo si $u_1 = u_2 = 0$, esto es, si $\mathbf{u} = \mathbf{0}$. Asimismo, $\langle \mathbf{0}, \mathbf{0} \rangle = 0$. Por lo tanto, (1) define un producto interior en $\mathbb{R}^2$.

Se pueden definir productos interiores semejantes a (1) en $\mathbb{R}^n$; los cuales surgen de manera natural en relación con "problemas de mínimos cuadrados ponderados", donde se asignan pesos a las diversas entradas de la suma para el producto interior, de modo que se dé mayor importancia a las medidas más confiables.

A partir de ahora, cuando en un espacio con producto interior intervengan polinomios u otras funciones, se escribirán las funciones de la manera acostumbrada, en lugar de usar el tipo de letra en negritas para identificar los vectores. Sin embargo, es importante recordar que cada función *es* un vector cuando se trata del elemento de un espacio vectorial.

EJEMPLO 2 Sean $t_0, \ldots, t_n$ números reales distintos. Para p y q en $\mathbb{P}_u$, defina

$$\langle p, q\rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n) \tag{2}$$

Los axiomas 1, 2 y 3 del producto interior se comprueban fácilmente. Para el axioma 4, observe que

$$\langle p, p\rangle = [p(t_0)]^2 + [p(t_1)]^2 + \cdots + [p(t_n)]^2 \geq 0$$

También, $\langle \mathbf{0}, \mathbf{0}\rangle = 0$. (Se seguirá usando un cero en negritas para identificar el polinomio cero, el vector cero en $\mathbb{P}_n$). Si $\langle p, p\rangle = 0$, entonces p debe desaparecer en $n + 1$ puntos: $t_0, \ldots, t_n$. Esto sólo es posible si p es el polinomio cero, porque el grado de p es menor que $n + 1$. Entonces (2) define a un producto interior en $\mathbb{P}_n$.

EJEMPLO 3 Sea V $\mathbb{P}_2$, con el producto interior del ejemplo 2, donde $t_0 = 0$, $t_1 = \frac{1}{2}$, y $t_2 = 1$. Sean $p(t) = 12t^2$ y $q(t) = 2t - 1$. Calcule $\langle p, q\rangle$ y $\langle q, q\rangle$.

Solución

$$\begin{aligned}\langle p, q\rangle &= p(0)q(0) + p\left(\tfrac{1}{2}\right)q\left(\tfrac{1}{2}\right) + p(1)q(1)\\ &= (0)(-1) + (3)(0) + (12)(1) = 12\\ \langle q, q\rangle &= [q(0)]^2 + [q\left(\tfrac{1}{2}\right)]^2 + [q(1)]^2\\ &= (-1)^2 + (0)^2 + (1)^2 = 2\end{aligned}$$

Longitudes, distancias y ortogonalidad

Sea V un espacio con producto interior, con el producto interior denotado mediante $\langle \mathbf{u}, \mathbf{v}\rangle$. Igual que en $\mathbb{R}^n$, la **longitud** o **norma** de un vector $\mathbf{v}$ se define como el escalar

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$$

De manera equivalente, $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v}\rangle$. (Esta definición tiene sentido, porque $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$, pero *no* establece que $\langle \mathbf{v}, \mathbf{v}\rangle$ sea una "suma de cuadrados", porque $\mathbf{v}$ debe ser un elemento de $\mathbb{R}^n$.)

Un **vector unitario** es aquel cuya longitud mide 1. La **distancia entre u y v** es $\|\mathbf{u} - \mathbf{v}\|$. Los vectores $\mathbf{u}$ y $\mathbf{v}$ son **ortogonales** si $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

EJEMPLO 4 Sea $\mathbb{P}_2$ tal que tenga el producto interior (2) del ejemplo 3. Encuentre las longitudes de los vectores $p(t) = 12t^2$ y $q(t) = 2t - 1$.

Solución

$$\begin{aligned}\|p\|^2 &= \langle p, p\rangle = [p(0)]^2 + \left[p\left(\tfrac{1}{2}\right)\right]^2 + [p(1)]^2\\ &= 0 + [3]^2 + [12]^2 = 153\\ \|p\| &= \sqrt{153}\end{aligned}$$

En el ejemplo 3, $\langle q, q\rangle = 2$. Entonces $\|q\| = \sqrt{2}$.

El proceso Gram-Schmidt

La existencia de bases ortogonales para subespacios de dimensión finita de un espacio con producto interior puede establecerse por medio del proceso Gram-Schmidt, de igual forma que en $\mathbb{R}^n$. Al aplicar este proceso, es posible plantear ciertas bases ortogonales que surgen con frecuencia en las aplicaciones.

La proyección ortogonal de un vector sobre un subespacio W con base ortogonal puede construirse como de costumbre. La proyección no depende de la selección de la base ortogonal y tiene las propiedades descritas en el teorema de la descomposición ortogonal y en el teorema de la mejor aproximación.

EJEMPLO 5 Sea V en $\mathbb{P}_4$ con el producto interior del ejemplo 2, que implica la evaluación de polinomios en $-2, -1, 0, 1$ y 2, y tome a $\mathbb{P}_2$ como un subespacio de V. Produzca una base ortogonal para $\mathbb{P}_2$ aplicando el proceso Gram-Schmidt a los polinomios 1, t y t^2.

Solución El producto interior depende sólo de los valores de un polinomio en -2, . . . , 2, así que se enlistan los valores de cada polinomio como un vector en $\mathbb{R}^5$, bajo el nombre del polinomio:[1]

$$\begin{array}{lccccc} \text{Polinomio:} & 1 & & t & & t^2 \\ \text{Vector de valores:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} & , & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} & , & \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix} \end{array}$$

El producto interior de dos polinomios en V es igual al producto interior (estándar) de sus vectores correspondientes en $\mathbb{R}^5$. Observe que t es ortogonal a la función constante 1. Así que tome a $p_0(t) = 1$ y $p_1(t) = t$. Para p_2, use los vectores en $\mathbb{R}^5$ para calcular la proyección de t^2 sobre Gen$\{p_0, p_1\}$:

$$\begin{aligned} \langle t^2, p_0 \rangle &= \langle t^2, 1 \rangle = 4 + 1 + 0 + 1 + 4 = 10 \\ \langle p_0, p_0 \rangle &= 5 \\ \langle t^2, p_1 \rangle &= \langle t^2, t \rangle = -8 + (-1) + 0 + 1 + 8 = 0 \end{aligned}$$

La proyección ortogonal de t^2 sobre Gen$\{1, t\}$ es $\frac{10}{5} p_0 + 0 p_1$. Así que

$$p_2(t) = t^2 - 2p_0(t) = t^2 - 2$$

[1]Cada polinomio en $\mathbb{P}_4$ está determinado de manera única por su valor en los cinco números -2, . . . , 2. De hecho, la correspondencia entre p y su vector de valores es un isomorfismo, es decir, una correspondencia uno a uno sobre $\mathbb{R}^5$ que conserva las combinaciones lineales.

Una base ortogonal para el subespacio $\mathbb{P}_2$ de V es:

$$\begin{array}{lccc} \text{Polinomio} & p_0 & p_1 & p_2 \\ \text{Vector de valores:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix} \end{array} \tag{3}$$

La mejor aproximación en espacios con producto interior

Un problema común en matemáticas aplicadas involucra un espacio vectorial V cuyos elementos son funciones. El problema consiste en aproximar una función f en V con una función g de un subespacio específico W de V. Lo "cercano" de la aproximación de f depende de la manera en que se defina $\|f - g\|$. Se considerará únicamente el caso en que la distancia entre f y g esté determinada por un producto interior. En este caso, la *mejor aproximación a f con funciones en W* es la proyección ortogonal de f sobre el subespacio W.

EJEMPLO 6 Sea V en $\mathbb{P}_4$ con el producto interior del ejemplo 5, y sean p_0, p_1 y p_2 la base ortogonal encontrada en el ejemplo 5 para el subespacio $\mathbb{P}_2$. Encuentre la aproximación óptima a $p(t) = 5 - \frac{1}{2}t^4$ mediante polinomios en $\mathbb{P}_2$.

Solución Los valores de p_0, p_1 y p_2 en los números $-2, -1, 0, 1$ y 2 se enumeran en vectores de $\mathbb{R}^5$ en la ecuación (3) anterior. Los valores correspondientes para p son -3, $9/2$, 5, $9/2$ y -3. Se calcula

$$\begin{array}{lll} \langle p, p_0 \rangle = 8, & \langle p, p_1 \rangle = 0, & \langle p, p_2 \rangle = -31 \\ \langle p_0, p_0 \rangle = 5, & & \langle p_2, p_2 \rangle = 14 \end{array}$$

Entonces la mejor aproximación a p en V por medio de polinomios en $\mathbb{P}_2$ es

$$\begin{aligned} \hat{p} = \text{proy}_{\mathbb{P}_2}\, p &= \frac{\langle p, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 + \frac{\langle p, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 + \frac{\langle p, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 \\ &= \tfrac{8}{5} p_0 + \tfrac{-31}{14} p_2 = \tfrac{8}{5} - \tfrac{31}{14}(t^2 - 2). \end{aligned}$$

Este polinomio es el más cercano a p de todos los polinomios en $\mathbb{P}_2$, cuando la distancia entre los polinomios se mide únicamente en $-2, -1, 0, 1$ y 2. Vea la figura 1 de la página 432.

Los polinomios p_0, p_1 y p_2 de los ejemplos 5 y 6 pertenecen a una clase de polinomios que en estadística se denominan *polinomios ortogonales*.[2] La ortogonalidad se refiere al tipo de producto interior descrito en el ejemplo 2.

[2]Vea *Statistics and Experimental Design in Engineering and the Physical Sciences*, de Norman L. Johnson y Fred C. Leone (Nueva York: John Wiley & Sons, 1964), págs. 424-436. Las tablas incluidas en las páginas 430 y 431 de esta fuente enumeran "polinomios ortogonales", que son simplemente los valores de los polinomios en números tales como $-2, -1, 0, 1$ y 2.

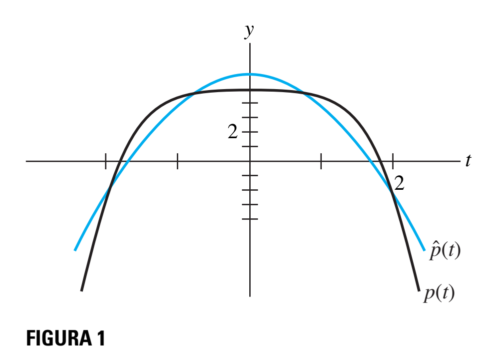


FIGURA 1

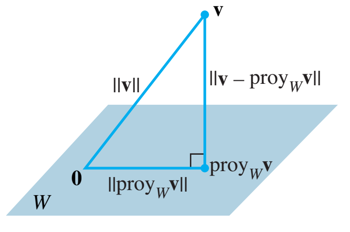


FIGURA 2
La hipotenusa es el lado más largo.

Dos desigualdades

Dado un vector **v** en un espacio con producto interior V y dado un subespacio de dimensión finita, puede aplicarse el teorema de Pitágoras a la descomposición ortogonal de **v** con respecto a W y obtener

$$\|\mathbf{v}\|^2 = \|\operatorname{proy}_W \mathbf{v}\|^2 + \|\mathbf{v} - \operatorname{proy}_W \mathbf{v}\|^2$$

Vea la figura 2. En particular, esto muestra que la norma de la proyección **v** sobre W no excede a la propia norma de **v**. Esta simple observación conduce a la siguiente importante desigualdad.

TEOREMA 16 **La desigualdad de Cauchy-Schwarz**

Para todas **u**, **v** en V,

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\| \tag{4}$$

DEMOSTRACIÓN Si $\mathbf{u} = \mathbf{0}$, entonces ambos lados de (4) son cero y, por lo tanto, en este caso (4) es cierta. (Vea el problema de práctica 1.) Si $\mathbf{u} \neq \mathbf{0}$, sea W el subespacio generado por **u**. Recuerde que $\|c\mathbf{u}\| = |c|\|\mathbf{u}\|$ para cualquier escalar c. Entonces

$$\|\operatorname{proy}_W \mathbf{v}\| = \left\| \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u} \right\| = \frac{|\langle \mathbf{v}, \mathbf{u}\rangle|}{|\langle \mathbf{u}, \mathbf{u}\rangle|}\|\mathbf{u}\| = \frac{|\langle \mathbf{v}, \mathbf{u}\rangle|}{\|\mathbf{u}\|^2}\|\mathbf{u}\| = \frac{|\langle \mathbf{u}, \mathbf{v}\rangle|}{\|\mathbf{u}\|}$$

Puesto que $\|\operatorname{proy}_W \mathbf{v}\| \leq \|\mathbf{v}$, se tiene que $\dfrac{|\langle \mathbf{u}, \mathbf{v}\rangle|}{\|\mathbf{u}\|} \leq \|\mathbf{v}\|$, de la cual se obtiene (4). ■

La desigualdad de Cauchy-Schwarz es útil en muchas ramas de las matemáticas. En los ejercicios se dan algunas aplicaciones sencillas. Aquí se necesita, principalmente, para demostrar otra desigualdad fundamental relacionada con las normas de los vectores. Vea la figura 3.

TEOREMA 17 **La desigualdad triangular**

Para todas $\mathbf{u}$, $\mathbf{v}$ en V,

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

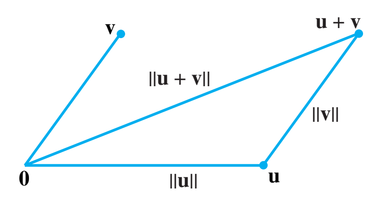


FIGURA 3
La longitud de los lados de un triángulo.

DEMOSTRACIÓN

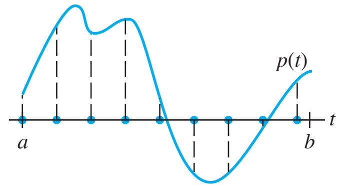
$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\ &\leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v}\rangle| + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 \qquad \text{Cauchy–Schwarz} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2\end{aligned}$$

Inmediatamente se deduce la desigualdad triangular al obtener la raíz cuadrada de ambos miembros. ■

Un producto interior para $C[a, b]$ (Se requiere cálculo)

Probablemente el espacio con producto interior más ampliamente usado en las aplicaciones sea el espacio vectorial $C[a, b]$ de todas las funciones continuas incluidas en un intervalo $a \leq t \leq b$, con un producto interior a describir enseguida.

Se inicia considerando un polinomio p y cualquier entero n mayor o igual al grado de p. Entonces p está en $\mathbb{P}_n$, y puede calcularse una "longitud" para p usando el producto interior del ejemplo 2 que implica la evaluación en $n + 1$ puntos de $[a, b]$. Sin embargo, esta longitud de p solamente capta el comportamiento en esos $n + 1$ puntos. Como p está en $\mathbb{P}_n$ para toda n grande, puede utilizarse una n mucho mayor, con muchos más puntos para el producto interior de "evaluación". Vea la figura 4.



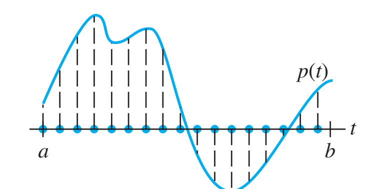
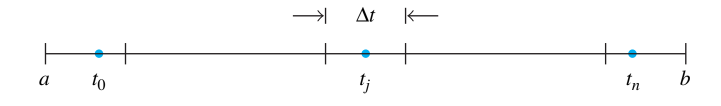


FIGURA 4 Uso de diferentes puntos de evaluación en $[a, b]$ para calcular $\|p\|^2$.

Se dividirá $[a, b]$ en $n + 1$ subintervalos de longitud $\Delta t = (b - a)/(n + 1)$, y sean $t_0, \ldots, t_n$ puntos arbitrarios en estos subintervalos.



Si n es grande, el producto interior en $\mathbb{P}_n$ determinado mediante $t_0, \ldots, t_n$ presentará una tendencia a dar un valor grande para $\langle p, p\rangle$, así que se reduce a escala y se divide entre

$n + 1$. Observe que $1/(n+1) = \Delta t/(b-a)$ y defina

$$\langle p, q\rangle = \frac{1}{n+1}\sum_{j=0}^{n} p(t_j)q(t_j) = \frac{1}{b-a}\left[\sum_{j=0}^{n} p(t_j)q(t_j)\Delta t\right]$$

Ahora, permita que n crezca en forma ilimitada. Puesto que los polinomios p y q son funciones continuas, la expresión incluida entre corchetes es una suma de Riemann que se aproxima a una integral definida y lleva a considerar el *valor promedio de* $p(t)q(t)$ sobre el intervalo $[a, b]$:

$$\frac{1}{b-a}\int_a^b p(t)q(t)\,dt$$

Esta cantidad está definida para polinomios de cualquier grado (de hecho, para todas las funciones continuas) y tiene todas las propiedades de un producto interior, como lo muestra el siguiente ejemplo. El factor de escala $1/(b-a)$ no es esencial, y a menudo se omite en aras de la simplicidad.

EJEMPLO 7 Para f, g de $C[a, b]$, sea

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,dt \tag{5}$$

Muestre que (5) define un producto interior en $C[a, b]$.

Solución Los axiomas 1, 2 y 3 del producto interior se derivan de las propiedades elementales de integrales definidas. Para el axioma 4, observe que

$$\langle f, f\rangle = \int_a^b [f(t)]^2\,dt \geq 0$$

La función $[f(t)]^2$ es continua y no negativa en $[a, b]$. Si la integral definida de $[f(t)]^2$ es cero, entonces $[f(t)]^2$ debe ser idénticamente cero en $[a, b]$, de acuerdo con un teorema de cálculo avanzado, en cuyo caso f es la función cero. Entonces $\langle f, f\rangle = 0$ implica que f es la función cero en $[a, b]$. Así que (5) define un producto interior en $C[a, b]$. ■

EJEMPLO 8 Sea V el espacio $C[0, 1]$ con el producto interior del ejemplo 7, y sea W el subespacio generado por los polinomios $p_1(t) = 1$, $p_2(t) = 2t - 1$, y $p_3(t) = 12t^2$. Use el proceso Gram-Schmidt para encontrar una base ortogonal para W.

Solución Sea $q_1 = p_1$, y calcule

$$\langle p_2, q_1\rangle = \int_0^1 (2t-1)(1)\,dt = (t^2 - t)\Big|_0^1 = 0$$

Entonces p_2 ya es ortogonal a q_1, y puede tomarse $q_2 = p_2$. Para la proyección de p_3 sobre $W_2 = \text{Gen}\{q_1, q_2\}$, se calcula

$$\langle p_3, q_1\rangle = \int_0^1 12t^2\cdot 1\,dt = 4t^3\Big|_0^1 = 4$$

$$\langle q_1, q_1\rangle = \int_0^1 1\cdot 1\,dt = t\Big|_0^1 = 1$$

$$\langle p_3, q_2 \rangle = \int_0^1 12t^2(2t-1)\,dt = \int_0^1 (24t^3 - 12t^2)\,dt = 2$$

$$\langle q_2, q_2 \rangle = \int_0^1 (2t-1)^2\,dt = \frac{1}{6}(2t-1)^3\Big|_0^1 = \frac{1}{3}$$

Entonces

$$\text{proy}_{W_2}\, p_3 = \frac{\langle p_3, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 + \frac{\langle p_3, q_2 \rangle}{\langle q_2, q_2 \rangle} q_2 = \frac{4}{1} q_1 + \frac{2}{1/3} q_2 = 4q_1 + 6q_2$$

y

$$q_3 = p_3 - \text{proy}_{W_2}\, p_3 = p_3 - 4q_1 - 6q_2$$

Como función, $q_3(t) = 12t^2 - 4 - 6(2t-1) = 12t^2 - 12t + 2$. La base ortogonal para el subespacio W es $\{q_1, q_2, q_3\}$.

PROBLEMAS DE PRÁCTICA

Use los axiomas del producto interior para verificar los siguientes enunciados.

1. $\langle \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle = 0$
2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$

6.7 EJERCICIOS

1. Sea $\mathbb{R}^2$ con el producto interior del ejemplo 1, y sean $\mathbf{x} = (1, 1)$ y $\mathbf{y} = (5, -1)$.
   a. Encuentre $\|\mathbf{x}\|$, $\|\mathbf{y}\|$ y $|\langle \mathbf{x}, \mathbf{y} \rangle|^2$.
   b. Describa todos los vectores (z_1, z_2) que sean ortogonales a $\mathbf{y}$.
2. Sea $\mathbb{R}^2$ con el producto interior del ejemplo 1. Muestre que la desigualdad Cauchy-Schwarz es válida para $\mathbf{x} = (3, -2)$ y $\mathbf{y} = (-2, 1)$. [*Sugerencia:* Estudie $|\langle \mathbf{x}, \mathbf{y} \rangle|^2$.]

Los ejercicios 3 a 8 se refieren a $\mathbb{P}_2$, con el producto interior dado por evaluación en -1, 0 y 1. (Vea el ejemplo 2.)

3. Calcule $\langle p, q \rangle$, donde $p(t) = 4 + t$, $q(t) = 5 - 4t^2$.
4. Calcule $\langle p, q \rangle$, donde $p(t) = 3t - t^2$, $q(t) = 3 - 2t^2$.
5. Calcule $\|p\|$ y $\|q\|$, para las p y q del ejercicio 3.
6. Calcule $\|p\|$ y $\|q\|$, para las p y q del ejercicio 4.
7. Determine la proyección ortogonal de q sobre el subespacio generado por p, para las p y q del ejercicio 3.
8. Determine la proyección ortogonal de q sobre el subespacio generado por p, para las p y q del ejercicio 4.
9. Sea $\mathbb{P}_3$ con el producto interior dado por evaluación en -3, -1, 1 y 3. Sean $p_0(t) = 1$, $p_1(t) = t$, y $p_2(t) = t^2$.
   a. Determine la proyección ortogonal de p_2 sobre el subespacio generado por p_0 y p_1.
   b. Encuentre un polinomio q ortogonal a p_0 y p_1, tal que $\{p_0, p_1, q\}$ sea una base ortogonal para Gen$\{p_0, p_1, p_2\}$. Escale al polinomio q de manera que su vector de valores en $(-3, -1, 1, 3)$ sea $(1, -1, -1, 1)$.
10. Sea $\mathbb{P}_3$ con el producto interior como en el ejercicio 9, siendo p_0, p_1 y q los polinomios allí descritos. Encuentre la aproximación óptima a $p(t) = t^3$ con polinomios en Gen$\{p_0, p_1, q\}$.
11. Sean p_0, p_1, p_2 los polinomios ortogonales descritos en el ejemplo 5, donde el producto interior en $\mathbb{P}_4$ está dado por la evaluación en -2, -1, 0, 1 y 2. Encuentre la proyección ortogonal de t^3 sobre Gen$\{p_0, p_1, p_2\}$.
12. Encuentre un polinomio p_3 tal que $\{p_0, p_1, p_2, p_3\}$ (vea el ejercicio 11) sea una base ortogonal para el subespacio $\mathbb{P}_3$ de $\mathbb{P}_4$. Escale el polinomio p_3 de manera que su vector de valores sea $(-1, 2, 0, -2, 1)$.

13. Sea A cualquier matriz invertible de $n \times n$. Muestre que para $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^n$, la fórmula $\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u}) \cdot (A\mathbf{v}) = (A\mathbf{u})^T(A\mathbf{v})$ define un producto interior en $\mathbb{R}^n$.

14. Sea T una transformación lineal uno a uno de un espacio vectorial V en $\mathbb{R}^n$. Muestre que para $\mathbf{u}$, $\mathbf{v}$ en V, la fórmula $\langle \mathbf{u}, \mathbf{v} \rangle = T(\mathbf{u}) \cdot T(\mathbf{v})$ define un producto interior en V.

Utilice los axiomas del producto interior y otros resultados de esta sección para verificar los enunciados de los ejercicios 15 a 18.

15. $\langle \mathbf{u}, c\mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ para todo escalar c.

16. Si $\langle \mathbf{u}, \mathbf{v} \rangle$ es un conjunto ortonormal en V. Entonces, $\|\mathbf{u} - \mathbf{v}\| = \sqrt{2}$.

17. $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$.

18. $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$.

19. Dados $a \geq 0$ y $b \geq 0$, sean $\mathbf{u} = \begin{bmatrix} \sqrt{a} \\ \sqrt{b} \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} \sqrt{b} \\ \sqrt{a} \end{bmatrix}$. Use la desigualdad de Cauchy-Schwarz para comparar la media geométrica $\sqrt{ab}$ con la media aritmética $(a + b)/2$.

20. Sean $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Use la desigualdad de Cauchy-Schwarz para mostrar que

$$\left(\frac{a + b}{2}\right)^2 \leq \frac{a^2 + b^2}{2}$$

Los ejercicios 21 a 24 se refieren a $V = C[0, 1]$, con el producto interior dado por una integral, como en el ejemplo 7.

21. Calcule $\langle f, g \rangle$, donde $f(t) = 1 - 3t^2$ y $g(t) = t - t^3$.

22. Calcule $\langle f, g \rangle$, donde $f(t) = 5t - 3$ y $g(t) = t^3 - t^2$.

23. Calcule $\|f\|$ para la f del ejercicio 21.

24. Calcule $\|g\|$ para la g del ejercicio 22.

25. Sea V el espacio $C[-1, 1]$ con el producto interior del ejemplo 7. Encuentre una base ortogonal para el subespacio generado por los polinomios 1, t y t^2. Los polinomios incluidos en esta base se llaman *polinomios de Legendre.*

26. Sea V el espacio $C[-2, 2]$ con el producto interior del ejemplo 7. Encuentre una base ortogonal para el subespacio generado por los polinomios 1, t y t^2.

27. **[M]** Sea $\mathbb{P}_4$ con el producto interior como el del ejemplo 5, y sean p_0, p_1, p_2 los polinomios ortogonales de ese ejemplo. Use un programa de matrices para aplicar el proceso de Gram-Schmidt al conjunto $\{p_0, p_1, p_2, t^3, t^4\}$ y crear una base ortogonal para $\mathbb{P}_4$.

28. **[M]** Sea V el espacio $C[0, 2\pi]$ con el producto interior del ejemplo 7. Aplique el proceso Gram-Schmidt y estructure una base ortogonal para el subespacio generado por $\{1, \cos t, \cos^2 t, \cos^3 t\}$. Use un programa de matrices o de computadora para calcular las integrales definidas apropiadas.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. De acuerdo con el axioma 1, $\langle \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle$. Entonces $\langle \mathbf{0}, \mathbf{v} \rangle = \langle 0\mathbf{v}, \mathbf{v} \rangle = 0\langle \mathbf{v}, \mathbf{v} \rangle$, según el axioma 3, así que $\langle \mathbf{0}, \mathbf{v} \rangle = 0$.
2. De acuerdo con los axiomas 1, 2, y de nuevo por el axioma 1, $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.

6.8 APLICACIONES DE LOS ESPACIOS CON PRODUCTO INTERIOR

Los ejemplos de esta sección sugieren cómo se presentan los espacios con producto interior definidos en la sección 6.7 en los problemas prácticos. El primer ejemplo está asociado con el masivo problema de mínimos cuadrados de actualizar el Nivel de Referencia Norteamericano, descrito en el ejemplo introductorio de este capítulo.

Mínimos cuadrados ponderados

Sea $\mathbf{y}$ un vector de n observaciones, $y_1, \ldots, y_n$, y suponga que se desea aproximar $\mathbf{y}$ con un vector $\hat{\mathbf{y}}$ perteneciente a algún subespacio específico de $\mathbb{R}^n$. (En la sección 6.5, $\hat{\mathbf{y}}$

se escribía como $A\mathbf{x}$ de manera que estaba en el espacio de columnas de A.) Denote las entradas de $\hat{\mathbf{y}}$ mediante $\hat{y}_1, \ldots, \hat{y}_n$. Entonces la *suma de los cuadrados del término de error*, o SS(E), al aproximar $\mathbf{y}$ con $\hat{\mathbf{y}}$ es

$$\text{SS(E)} = (y_1 - \hat{y}_1)^2 + \cdots + (y_n - \hat{y}_n)^2 \tag{1}$$

Esto simplemente es $\|\mathbf{y} - \hat{\mathbf{y}}\|^2$, usando la longitud estándar en $\mathbb{R}^n$.

Ahora suponga que las mediciones que produjeron las entradas de $\mathbf{y}$ no son todas igualmente confiables. (Éste fue el caso para el Nivel de Referencia Norteamericano, puesto que las mediciones se tomaron durante un periodo de 140 años. Como ejemplo adicional, las entradas de $\mathbf{y}$ podrían haberse calculado a partir de varias muestras de mediciones, con muestras de tamaños desiguales.) Entonces resulta adecuado ponderar los errores cuadrados de (1) de manera que se dé más importancia a las mediciones más confiables.[1] Al denotar los pesos mediante $w_1^2, \ldots, w_n^2$, entonces la suma ponderada de los cuadrados para el error es

$$\text{SS(E) Ponderada} = w_1^2(y_1 - \hat{y}_1)^2 + \cdots + w_n^2(y_n - \hat{y}_n)^2 \tag{2}$$

Esto es el cuadrado de la longitud de $\mathbf{y} - \hat{\mathbf{y}}$, donde la longitud se obtiene a partir de un producto interior análogo al del ejemplo 1 de la sección 6.7, a saber,

$$\langle \mathbf{x}, \mathbf{y} \rangle = w_1^2 x_1 y_1 + \cdots + w_n^2 x_n y_n$$

Algunas veces es conveniente transformar un problema de mínimos cuadrados ponderados en un problema ordinario de mínimos cuadrados equivalente. Sea W la matriz diagonal con $w_1, \ldots, w_n$ (positivos) en su diagonal, de modo que

$$W\mathbf{y} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & w_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} w_1 y_1 \\ w_2 y_2 \\ \vdots \\ w_n y_n \end{bmatrix}$$

con una expresión similar para $W\hat{\mathbf{y}}$. Observe que el término j-ésimo de (2) puede escribirse como

$$w_j^2(y_j - \hat{y}_j)^2 = (w_j y_j - w_j \hat{y}_j)^2$$

Se deduce que la SS(E) ponderada de (2) es el cuadrado de la longitud ordinaria en $\mathbb{R}^n$ de $W\mathbf{y} - W\hat{\mathbf{y}}$, lo cual se escribe como $\|W\mathbf{y} - W\hat{\mathbf{y}}\|^2$.

Ahora suponga que el vector de aproximación $\hat{\mathbf{y}}$ va a ser estructurado como las columnas de una matriz A. Entonces se busca un $\hat{\mathbf{x}}$ que haga a $A\hat{\mathbf{x}} = \hat{\mathbf{y}}$ tan cercano como sea posible. Sin embargo, la medida de cercanía es el error ponderado.

$$\|W\mathbf{y} - W\hat{\mathbf{y}}\|^2 = \|W\mathbf{y} - WA\hat{\mathbf{x}}\|^2$$

Entonces $\hat{\mathbf{x}}$ es la solución (ordinaria) por mínimos cuadrados de la ecuación

$$WA\mathbf{x} = W\mathbf{y}$$

[1]Nota para los lectores con conocimientos de estadística: Suponga que los errores al medir las y_i son variables aleatorias independientes con medias iguales a cero y varianzas de $\sigma_1^2, \ldots, \sigma_n^2$. Entonces los pesos apropiados en (2) son $w_i^2 = 1/\sigma_i^2$. A mayor varianza del error, menor peso.

La ecuación normal para la solución por mínimos cuadrados es

$$(WA)^T WA\mathbf{x} = (WA)^T W\mathbf{y}$$

EJEMPLO 1 Encuentre la línea de mínimos cuadrados $y = \beta_0 + \beta_1 x$ que mejor se ajuste a los datos $(-2, 3)$, $(-1, 5)$, $(0, 5)$, $(1, 4)$, $(2, 3)$. Suponga que los errores al medir los valores de y de los dos últimos puntos de datos son mayores que para los otros puntos. Pondere estos datos en la mitad de lo que lo haría con el resto de los datos.

Solución De igual forma que en la sección 6.6, escriba X en lugar de la matriz A y $\boldsymbol{\beta}$ para el vector $\mathbf{x}$, y obtenga

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix}$$

Para una matriz de ponderación, elija W con entradas diagonales 2, 2, 2, 1 y 1. Multiplicando a la izquierda por W, se escalan las filas de X y $\mathbf{y}$:

$$WX = \begin{bmatrix} 2 & -4 \\ 2 & -2 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad W\mathbf{y} = \begin{bmatrix} 6 \\ 10 \\ 10 \\ 4 \\ 3 \end{bmatrix}$$

Para la ecuación normal, calcule

$$(WX)^T WX = \begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \quad \text{y} \quad (WX)^T W\mathbf{y} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

y resuelva

$$\begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

La solución de la ecuación normal es (con dos dígitos significativos) $\beta_0 = 4.3$ y $\beta_1 = .20$. La línea deseada es

$$y = 4.3 + .20x$$

En contraste, la línea de mínimos cuadrados ordinaria para estos datos es

$$y = 4.0 - .10x$$

Ambas líneas se muestran en la figura 1.

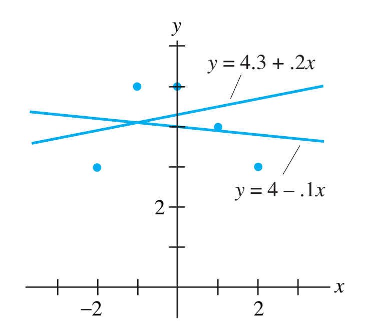


FIGURA 1 Líneas de mínimos cuadrados ponderada y ordinaria.

Análisis de la tendencia de los datos

Sea f tal que represente una función desconocida cuyos valores se conocen (quizá sólo aproximadamente) en $t_0, \ldots, t_n$. Si hay una "tendencia lineal" en los datos $f(t_0), \ldots,$

$f(t_n)$, entonces podría esperarse aproximar los valores de f mediante una función de la forma $\beta_0 + \beta_1 t$. Si los datos tienen una "tendencia cuadrática", se intentaría con una función de la forma $\beta_0 + \beta_1 t + \beta_2 t^2$. Esto se explicó en la sección 6.6, desde un punto de vista diferente.

En algunos problemas de estadística, es importante poder separar la tendencia lineal de la tendencia cuadrática (y posiblemente de tendencias cúbicas o de mayor orden). Por ejemplo, suponga que ciertos ingenieros están analizando el desempeño de un nuevo automóvil y que $f(t)$ representa la distancia, en el tiempo t, medida entre el automóvil y algún punto de referencia. Si el automóvil viaja con velocidad constante, entonces la gráfica de $f(t)$ debería ser una línea recta cuya pendiente es la velocidad del automóvil. Si de pronto se presiona el pedal del acelerador hasta el fondo, la gráfica de $f(t)$ cambiará para incluir un término cuadrático, y posiblemente uno cúbico (debido a la aceleración). Para analizar la capacidad del automóvil de rebasar otro automóvil, por ejemplo, los ingenieros podrían querer separar los componentes cuadrático y cúbico del término lineal.

Si se aproxima la función empleando una curva de la forma $y = \beta_0 + \beta_1 t + \beta_2 t^2$, puede ser que el coeficiente β_2 no proporcione la información deseada acerca de la tendencia cuadrática de los datos, porque podría no ser "independiente" en sentido estadístico, de los otros β_i. Para realizar lo que se conoce como **análisis de tendencia** de los datos, se introduce un producto interior en el espacio $\mathbb{P}_n$ análogo al que se dio en el ejemplo 2 de la sección 6.7. Para p, q en $\mathbb{P}_n$, se define

$$\langle p, q\rangle = p(t_0)q(t_0) + \cdots + p(t_n)q(t_n)$$

En la práctica, los estadísticos rara vez necesitan considerar las tendencias en datos de grado mayor que cúbico o cuadrático. Así que sea p_0, p_1, p_2, p_3 una base ortogonal del subespacio $\mathbb{P}_3$ de $\mathbb{P}_n$, obtenida al aplicar el proceso Gram-Schmidt a los polinomios 1, t, t^2 y t^3. De acuerdo con el ejercicio suplementario 11 del capítulo 2, existe un polinomio g en $\mathbb{P}_n$ cuyos valores en $t_0, \ldots, t_n$ coinciden con los de la función f desconocida. Sea $\hat{g}$ la proyección ortogonal (con respecto al producto interior dado) de g sobre $\mathbb{P}_3$, es decir,

$$\hat{g} = c_0 p_0 + c_1 p_1 + c_2 p_2 + c_3 p_3$$

Entonces $\hat{g}$ se denomina **función de tendencia** cúbica, y $c_0, \ldots, c_3$ son los **coeficientes de tendencia** de los datos. El coeficiente c_1 mide la tendencia lineal, c_2 la tendencia cuadrática, y c_3 la tendencia cúbica. Resulta que si los datos tienen ciertas propiedades, estos coeficientes son estadísticamente independientes.

Como $p_0, \ldots, p_3$ son ortogonales, los coeficientes de tendencia pueden calcularse uno a la vez, independientemente uno del otro. (Recuerde que $c_i = \langle g, p_i\rangle / \langle p_i, p_i\rangle$.) Puede no considerarse p_3 y c_3 si únicamente se desea la tendencia cuadrática. Y si, por ejemplo, se necesitara determinar la tendencia a la cuarta, sería necesario encontrar (mediante Gram-Schmidt) únicamente un polinomio p_4 en $\mathbb{P}_4$ que sea ortogonal a $\mathbb{P}_3$ y calcular $\langle g, p_4\rangle / \langle p_4, p_4\rangle$.

EJEMPLO 2 El uso más común y sencillo del análisis de tendencia ocurre cuando los puntos $t_0, \ldots, t_n$ pueden ajustarse de manera que tengan una separación uniforme y sumen cero. Ajuste una función de tendencia cuadrática a los datos $(-2, 3)$, $(-1, 5)$, $(0, 5)$, $(1, 4)$ y $(2, 3)$.

Solución Se aplica una escala adecuada a las coordenadas t para usar los polinomios ortogonales encontrados en el ejemplo 5 de la sección 6.7. Se tiene

$$\begin{array}{lcccc} \text{Polinomio:} & p_0 & p_1 & p_2 & \text{Datos: } g \\ \text{Vector de valores:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix} \end{array}$$

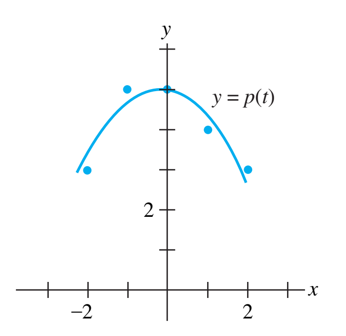


FIGURA 2
Aproximación mediante una función de tendencia cuadrática.

En los cálculos sólo intervienen estos vectores, no las fórmulas específicas para los polinomios ortogonales: la mejor aproximación a los datos con polinomios en $\mathbb{P}_2$ es la proyección ortogonal dada por

$$\begin{aligned} \hat{p} &= \frac{\langle g, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 + \frac{\langle g, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 + \frac{\langle g, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 \\ &= \tfrac{20}{5} p_0 - \tfrac{1}{10} p_1 - \tfrac{7}{14} p_2 \end{aligned}$$

y

$$\hat{p}(t) = 4 - .1t - .5(t^2 - 2) \tag{3}$$

Como el coeficiente de p_2 no es extremadamente pequeño, sería razonable concluir que la tendencia es, por lo menos, cuadrática. Esto se confirma con la gráfica de la figura 2.

Series de Fourier (se requiere cálculo)

A menudo las funciones continuas se aproximan mediante combinaciones lineales de funciones seno y coseno. Por ejemplo, una función continua podría representar una onda de sonido, una señal eléctrica de algún tipo, o el movimiento de un sistema mecánico que vibra.

En aras de la simplicidad, considere funciones en $0 \leq t \leq 2\pi$. Resulta que cualquier función en $C[0, 2\pi]$ puede aproximarse tanto como se desee mediante una función de la forma

$$\frac{a_0}{2} + a_1 \cos t + \cdots + a_n \cos nt + b_1 \operatorname{sen} t + \cdots + b_n \operatorname{sen} nt \tag{4}$$

para un valor de n lo suficientemente grande. La función (4) se llama **polinomio trigonométrico.** Si a_n y b_n no son ambos cero, se afirma que el polinomio es de **orden *n***. La conexión entre los polinomios trigonométricos y otras funciones de $C[0, 2\pi]$ depende del hecho de que para cualquier $n \geq 1$, el conjunto

$$\{1, \cos t, \cos 2t, \ldots, \cos nt, \operatorname{sen} t, \operatorname{sen} 2t, \ldots, \operatorname{sen} nt\} \tag{5}$$

sea ortogonal con respecto al producto interior

$$\langle f, g \rangle = \int_0^{2\pi} f(t) g(t)\, dt \tag{6}$$

Esta ortogonalidad se verifica tal como aparece en el siguiente ejemplo y en los ejercicios 5 y 6.

EJEMPLO 3 Sea $C[0, 2\pi]$ con el producto interior (6), y sean m y n enteros positivos diferentes. Muestre que $\cos mt$ y $\cos nt$ son ortogonales.

Solución Se utiliza una identidad trigonométrica. Cuando $m \neq n$,

$$\begin{aligned}\langle \cos mt, \cos nt \rangle &= \int_0^{2\pi} \cos mt \cos nt \, dt \\ &= \frac{1}{2}\int_0^{2\pi} [\cos(mt+nt) + \cos(mt-nt)]\, dt \\ &= \frac{1}{2}\left[\frac{\operatorname{sen}(mt+nt)}{m+n} + \frac{\operatorname{sen}(mt-nt)}{m-n}\right]\Bigg|_0^{2\pi} = 0\end{aligned}$$

Sea W el subespacio de $C[0, 2\pi]$ generado por las funciones de (5). Dada f en $C[0, 2\pi]$, la mejor aproximación a f con funciones en W se llama **aproximación de Fourier de orden *n*** a f en $[0, 2\pi]$. Como las funciones de (5) son ortogonales, la mejor aproximación está dada por la proyección ortogonal sobre W. En este caso, los coeficientes a_k y b_k de (4) se denominan **coeficientes de Fourier** de f. La fórmula estándar para la proyección ortogonal muestra que

$$a_k = \frac{\langle f, \cos kt \rangle}{\langle \cos kt, \cos kt \rangle}, \quad b_k = \frac{\langle f, \operatorname{sen} kt \rangle}{\langle \operatorname{sen} kt, \operatorname{sen} kt \rangle}, \quad k \geq 1$$

El ejercicio 7 solicita demostrar que $\langle \cos kt, \cos kt \rangle = \pi$ y que $\langle \operatorname{sen} kt, \operatorname{sen} kt \rangle = \pi$. Entonces

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(t) \cos kt \, dt, \quad b_k = \frac{1}{\pi}\int_0^{2\pi} f(t) \operatorname{sen} kt \, dt \tag{7}$$

El coeficiente de la función 1 (constante) en la proyección ortogonal es

$$\frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{2\pi}\int_0^{2\pi} f(t)\cdot 1 \, dt = \frac{1}{2}\left[\frac{1}{\pi}\int_0^{2\pi} f(t)\cos(0\cdot t)\, dt\right] = \frac{a_0}{2}$$

donde a_0 está definida por (7) para $k = 0$. Esto explica por qué el término constante (4) se escribe como $a_0/2$.

EJEMPLO 4 Encuentre la aproximación de Fourier de orden n a la función $f(t) = t$ en el intervalo $[0, 2\pi]$.

Solución Se calcula

$$\frac{a_0}{2} = \frac{1}{2}\cdot\frac{1}{\pi}\int_0^{2\pi} t \, dt = \frac{1}{2\pi}\left[\frac{1}{2}t^2\Big|_0^{2\pi}\right] = \pi$$

y para $k > 0$, utilizando integración por partes,

$$\begin{aligned}a_k &= \frac{1}{\pi}\int_0^{2\pi} t \cos kt \, dt = \frac{1}{\pi}\left[\frac{1}{k^2}\cos kt + \frac{t}{k}\operatorname{sen} kt\right]_0^{2\pi} = 0 \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} t \operatorname{sen} kt \, dt = \frac{1}{\pi}\left[\frac{1}{k^2}\operatorname{sen} kt - \frac{t}{k}\cos kt\right]_0^{2\pi} = -\frac{2}{k}\end{aligned}$$

Entonces la aproximación de Fourier de orden n a $f(t) = t$ es

$$\pi - 2\,\text{sen}\,t - \text{sen}\,2t - \frac{2}{3}\,\text{sen}\,3t - \cdots - \frac{2}{n}\,\text{sen}\,nt$$

En la figura 3 se muestran las aproximaciones de Fourier de tercer y cuarto orden a f.

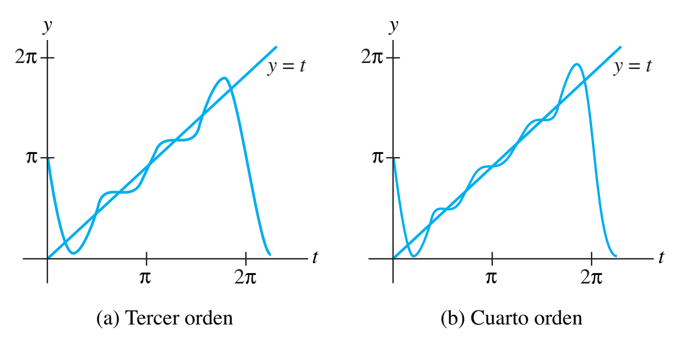


FIGURA 3 Aproximaciones de Fourier a la función $f(t) = t$.

La norma de la diferencia entre f y una aproximación de Fourier se llama **error cuadrado medio** de la aproximación. (El término *medio* se refiere al hecho de que la norma está determinada por una integral.) Puede demostrarse que el error cuadrado medio se aproxima a cero cuando aumenta el orden de la aproximación de Fourier. Por esta razón, es común escribir

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty}(a_m \cos mt + b_m\,\text{sen}\,mt)$$

Esta expresión para $f(t)$ es la **serie de Fourier** para f en $[0, 2\pi]$. El término $a_m \cos mt$, por ejemplo, es la proyección de f sobre el subespacio unidimensional generado por $\cos mt$.

PROBLEMAS DE PRÁCTICA

1. Sean $q_1(t) = 1$, $q_2(t) = t$, y $q_3(t) = 3t^2 - 4$. Verifique si $\{q_1, q_2, q_3\}$ es un conjunto ortogonal en $C[-2, 2]$ con el producto interior del ejemplo 7 dado en la sección 6.7 (integración de -2 a 2).
2. Encuentre las aproximaciones de Fourier de primer y tercer orden a

$$f(t) = 3 - 2\,\text{sen}\,t + 5\,\text{sen}\,2t - 6\cos 2t$$

6.8 EJERCICIOS

1. Encuentre la línea de mínimos cuadrados $y = \beta_0 + \beta_1 x$ que se ajuste mejor a los datos (−2, 0), (−1, 0), (0, 2), (1, 4), (2, 4), suponiendo que los puntos de datos primero y último son menos confiables. Pondere éstos a la mitad de los tres puntos interiores.

2. Suponga que en un problema de mínimos cuadrados ponderados, 5 de 25 puntos de datos tienen una medición y que es menos confiable que las otras mediciones, y que se deben ponderar en la mitad de lo que se ponderan los otros 20 puntos. Un método apropiado para resolver esto consiste en ponderar los 20 puntos mediante un factor de 1 y los otros 5 mediante un factor de $\frac{1}{2}$. Un segundo método es ponderar los 20 puntos por un factor de 2 y los otros 5 por un factor de 1. ¿Producen diferentes resultados los dos métodos? Explique su respuesta.

3. Ajuste una función de tendencia cúbica a los datos del ejemplo 2. El polinomio cúbico ortogonal es $p_3(t) = \frac{5}{6}t^3 - \frac{17}{6}t$.

4. Para hacer un análisis de tendencia de seis puntos de datos espaciados regularmente, se pueden usar polinomios ortogonales con respecto a la evaluación en los puntos $t = -5, -3, -1, 1, 3$ y 5.

a. Muestre que los primeros tres polinomios ortogonales son

$$p_0(t) = 1, \quad p_1(t) = t, \quad \text{y} \quad p_2(t) = \tfrac{3}{8}t^2 - \tfrac{35}{8}$$

(El polinomio p_2 está escalado de manera que sus valores en los puntos de evaluación sean enteros pequeños.)

b. Ajuste una función de tendencia cuadrática a los datos

$$(-5, 1), (-3, 1), (-1, 4), (1, 4), (3, 6), (5, 8)$$

En los ejercicios 5 a 14, el espacio es $C[0, 2\pi]$ con el producto interior (6).

5. Muestre que sen mt y sen nt son ortogonales cuando $m \neq n$.

6. Muestre que sen mt y cos nt son ortogonales para todos los valores enteros positivos de m y n.

7. Muestre que $\|\cos kt\|^2 = \pi$ y $\|\text{sen}\, kt\|^2 = \pi$ para $k > 0$.

8. Encuentre la aproximación de Fourier de tercer orden a $f(t) = t - 1$.

9. Encuentre la aproximación de Fourier de tercer orden a $f(t) = 2\pi - t$.

10. Encuentre la aproximación de Fourier de tercer orden a la *función de onda cuadrada*, $f(t) = 1$ para $0 \leq t < \pi$ y $f(t) = -1$ para $\pi \leq t < 2\pi$.

11. Encuentre la aproximación de Fourier de tercer orden a $\text{sen}^2\, t$, sin realizar cálculos de integración.

12. Encuentre la aproximación de Fourier de tercer orden a $\cos^3 t$, sin calcular ninguna integral.

13. Explique por qué un coeficiente de Fourier de la suma de dos funciones es la suma de los coeficientes de Fourier correspondientes a las dos funciones.

14. Suponga que los primeros coeficientes de Fourier de alguna función f en $C[0, 2\pi]$ son a_0, a_1, a_2 y b_1, b_2, b_3. ¿Cuál de los siguientes polinomios trigonométricos es más cercano a f? Defienda su respuesta.

$$g(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \,\text{sen}\, t$$

$$h(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \,\text{sen}\, t + b_2 \,\text{sen}\, 2t$$

15. **[M]** En referencia a los datos del ejercicio 13 de la sección 6.6, relativos al desempeño de un avión durante el despegue. Suponga que los posibles errores de medición se vuelven mayores conforme aumenta la velocidad del avión, y sea W la matriz diagonal de pesos cuyas entradas son 1, 1, 1, .9, .9, .8, .7, .6, .5, .4, .3, .2 y .1. Encuentre la curva cúbica que se ajuste a los datos con el menor error de mínimos cuadrados ponderados, y utilícela para estimar la velocidad del avión cuando $t = 4.5$ segundos.

16. **[M]** Sean f_4 y f_5 las aproximaciones de Fourier de cuarto y quinto orden en $C[0, 2\pi]$ a la función de onda cuadrada del ejercicio 10. Trace gráficas separadas de f_4 y f_5 en el intervalo $[0, 2\pi]$, y produzca una gráfica de f_5, en $[-2\pi, 2\pi]$.

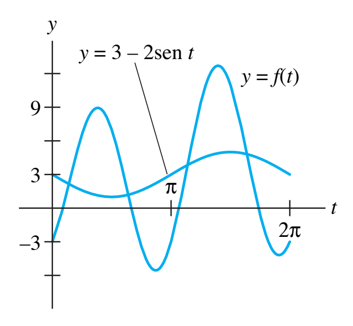
SG La linealidad de una proyección ortogonal 6 a 25 (The Linearity of an Orthogonal Projection 6-25)

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Calcule

$$\langle q_1, q_2 \rangle = \int_{-2}^{2} 1 \cdot t\, dt = \frac{1}{2}t^2 \Big|_{-2}^{2} = 0$$

$$\langle q_1, q_3 \rangle = \int_{-2}^{2} 1 \cdot (3t^2 - 4)\, dt = (t^3 - 4t) \Big|_{-2}^{2} = 0$$



Aproximaciones de primer y tercer orden a $f(t)$.

$$\langle q_2, q_3\rangle = \int_{-2}^{2} t\cdot(3t^2-4)\,dt = \left(\frac{3}{4}t^4 - 2t^2\right)\Big|_{-2}^{2} = 0$$

2. La aproximación de Fourier de tercer orden a f es la mejor aproximación en $C[0, 2\pi]$ a f con funciones (vectores) del subespacio generado por 1, cos t, cos $2t$, cos $3t$, sen t, sen $2t$ y sen $3t$. Pero desde luego que f está *en* este subespacio, así que f es su propia mejor aproximación:

$$f(t) = 3 - 2\,\text{sen}\,t + 5\,\text{sen}\,2t - 6\cos 2t$$

Para la aproximación de primer orden, la función más cercana a f en el subespacio $W = \text{Gen}\{1, \cos t, \text{sen}\,t\}$ es $3 - 2$ sen t. Los otros dos términos de la fórmula para $f(t)$ son ortogonales a las funciones en W, así que no contribuyen en nada a las integrales que proporcionan los coeficientes de Fourier para una aproximación de primer orden.

CAPÍTULO 6 EJERCICIOS SUPLEMENTARIOS

1. Los enunciados siguientes se refieren a vectores en $\mathbb{R}^n$ (o $\mathbb{R}^m$) con el producto interior estándar. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

a. La longitud de cualquier vector es un número positivo.

b. Un vector $\mathbf{v}$ y su negativo $-\mathbf{v}$ tienen la misma longitud.

c. La distancia entre $\mathbf{u}$ y $\mathbf{v}$ es $\|\mathbf{u} - \mathbf{v}\|$.

d. Si r es un escalar cualquiera, entonces $\|r\mathbf{v}\| = r\|\mathbf{v}\|$.

e. Si dos vectores son ortogonales, entonces son linealmente independientes.

f. Si $\mathbf{x}$ es ortogonal tanto a $\mathbf{u}$ como a $\mathbf{v}$, entonces $\mathbf{x}$ debe ser ortogonal a $\mathbf{u} - \mathbf{v}$.

g. Si $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

h. Si $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

i. La proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$ es un múltiplo escalar de $\mathbf{y}$.

j. Si un vector $\mathbf{y}$ coincide con su proyección ortogonal sobre un subespacio W, entonces $\mathbf{y}$ está en W.

k. El conjunto de todos los vectores en $\mathbb{R}^n$ que son ortogonales a un vector fijo es un subespacio de $\mathbb{R}^n$.

l. Si W es un subespacio de $\mathbb{R}^n$, entonces W y $W^\perp$ no tienen vectores en común.

m. Si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal, y si c_1, c_2 y c_3 son escalares, entonces $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2, c_3\mathbf{v}_3\}$ es un conjunto ortogonal.

n. Si una matriz U tiene columnas ortonormales, entonces $UU^T = I$.

o. Una matriz cuadrada con columnas ortogonales es una matriz ortogonal.

p. Si una matriz cuadrada tiene columnas ortonormales, entonces también tiene filas ortonormales.

q. Si W es un subespacio, entonces $\|\text{proy}_W\,\mathbf{v}\|^2 + \|\mathbf{v} - \text{proy}_W\,\mathbf{v}\|^2 = \|\mathbf{v}\|^2$.

r. Una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es el vector $A\hat{\mathbf{x}}$ en Col A más cercano a $\mathbf{b}$, tal que $\|\mathbf{b} - A\hat{\mathbf{x}}\| \le \|\mathbf{b} - A\mathbf{x}\|$ para toda $\mathbf{x}$.

s. Las ecuaciones normales para una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ están dadas por $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

2. Sea $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ un conjunto ortonormal. Verifique por inducción la siguiente igualdad, comenzando con $p = 2$. Si $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, entonces

$$\|\mathbf{x}\|^2 = \{c_1\}^2 + \cdots + \{c_p\}^2$$

3. Sea $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ un conjunto ortonormal en $\mathbb{R}^n$. Verifique la siguiente desigualdad, llamada *desigualdad de Bessel*, la cual es verdadera para cada $\mathbf{x}$ en $\mathbb{R}^n$:

$$\|\mathbf{x}\|^2 \ge \{\mathbf{x}\cdot\mathbf{v}_1\}^2 + \{\mathbf{x}\cdot\mathbf{v}_2\}^2 + \cdots + \{\mathbf{x}\cdot\mathbf{v}_p\}^2$$

4. Sea U una matriz ortogonal de $n \times n$. Muestre que si $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ es una base ortonormal para $\mathbb{R}^n$, entonces también lo es $\{U\mathbf{v}_1, \dots, U\mathbf{v}_n\}$.

5. Muestre que si una matriz U $n \times n$ satisface $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$ para toda $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$, entonces U es una matriz ortogonal.

6. Muestre que si U es una matriz ortogonal, entonces cualquier valor propio real de U debe ser ± 1.

7. Una *matriz de Householder*, o un *reflector elemental*, tiene la forma $Q = I - 2\mathbf{u}\mathbf{u}^T$, donde $\mathbf{u}$ es un vector unitario. (Vea el ejercicio 13 en los ejercicios suplementarios del capítulo 2.) Muestre que Q es una matriz ortogonal. Los reflectores elementales se usan con frecuencia en programas de cómputo para producir una factorización QR de una matriz A. Si

A tiene columnas linealmente independientes, entonces la multiplicación por la izquierda mediante una sucesión de reflectores elementales puede producir una matriz triangular superior.)

8. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal que conserva las longitudes; esto es, $\|T(\mathbf{x})\| = \|\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

a. Muestre que también T conserva la ortogonalidad; esto es $T(\mathbf{x}) \cdot T(\mathbf{y}) = 0$ siempre que $\mathbf{x} \cdot \mathbf{y} = 0$.

b. Muestre que la matriz estándar de T es una matriz ortogonal.

9. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores linealmente independientes en $\mathbb{R}^n$ que *no* sean ortogonales. Describa cómo encontrar la mejor aproximación a $\mathbf{z}$ en $\mathbb{R}^n$ mediante vectores de la forma $x_1\mathbf{u} + x_2\mathbf{v}$ sin crear primero una base ortogonal para Gen$\{\mathbf{u}, \mathbf{v}\}$.

10. Suponga que las columnas de A son linealmente independientes. Determine lo que sucede a la solución por mínimos cuadrados $\hat{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ cuando $\mathbf{b}$ se reemplaza por $c\mathbf{b}$ para algún escalar c distinto de cero.

11. Si a, b y c son números distintos, entonces el siguiente sistema es inconsistente porque las gráficas de las ecuaciones son planos paralelos. Muestre que el conjunto de todas las soluciones por mínimos cuadrados es precisamente el plano cuya ecuación es $x - 2y + 5z = (a + b + c)/3$.

$$\begin{aligned} x - 2y + 5z &= a \\ x - 2y + 5z &= b \\ x - 2y + 5z &= c \end{aligned}$$

12. Considere el problema de encontrar un valor propio de una matriz A de $n \times n$ cuando se conoce un vector propio aproximado $\mathbf{v}$. Como $\mathbf{v}$ no es exactamente correcto, la ecuación

$$A\mathbf{v} = \lambda\mathbf{v} \qquad (1)$$

probablemente no tendrá solución. Sin embargo, puede estimarse λ mediante una solución por mínimos cuadrados cuando (1) se ve de manera apropiada. Piense en $\mathbf{v}$ como una matriz V de $n \times 1$, piense en λ como un vector en $\mathbb{R}^1$, y denote $A\mathbf{v}$ con el símbolo $\mathbf{b}$. Entonces (1) se convierte en $\mathbf{b} = \lambda V$, que también puede escribirse como $V\lambda = \mathbf{b}$. Encuentre la solución por mínimos cuadrados de este sistema de n ecuaciones en la única incógnita λ, y escriba esta solución usando los símbolos originales. La estimación resultante para λ se llama *cociente de Rayleigh*. Vea los ejercicios 11 y 12 de la sección 5.8.

13. Siga los pasos que se dan más adelante para demostrar las siguientes relaciones entre los cuatro subespacios fundamentales de una matriz A de $m \times n$.

$$\text{Fil } A = (\text{Nul } A)^\perp, \quad \text{Col } A = (\text{Nul } A^T)^\perp$$

a. Muestre que Fil A está contenido en (Nul $A)^\perp$. (Muestre que si $\mathbf{x}$ está en Fil A, entonces $\mathbf{x}$ es ortogonal a toda $\mathbf{u}$ en Nul A.)

b. Suponga que rango $A = r$. Encuentre dim Nul A y dim(Nul $A)^\perp$, y luego deduzca de (a) que Fil A = (Nul $A)^\perp$. [*Sugerencia:* Estudie los ejercicios de la sección 6.3.]

c. Explique por qué Col A = (Nul $A^T)^\perp$.

14. Explique por qué una ecuación $A\mathbf{x} = \mathbf{b}$ tiene solución si, y sólo si, $\mathbf{b}$ es ortogonal a todas las soluciones de la ecuación $A^T\mathbf{x} = \mathbf{0}$.

Los ejercicios 15 y 16 están relacionados con la *factorización de Schur* (real) de una matriz A de $n \times n$ de la forma $A = URU^T$, donde U es una matriz ortogonal y R es una matriz triangular superior de $n \times$ n.[1]

15. Muestre que si A admite una factorización de Schur (real), $A = URU^T$, entonces tiene n valores propios reales, contando las multiplicidades.

16. Sea A una matriz de $n \times n$ con n valores propios reales, contando multiplicidades, denotados mediante $\lambda_1, \ldots, \lambda_n$. Puede mostrarse que A admite una factorización de Schur (real). Los incisos (a) y (b) presentan las ideas clave de la demostración. El resto de la demostración equivale a repetir (a) y (b) para matrices menores sucesivas y concatenar luego los resultados.

a. Sea $\mathbf{u}_1$ un vector propio unitario correspondiente a λ_1, sean $\mathbf{u}_2, \ldots, \mathbf{u}_n$ cualesquiera otros vectores tales que $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ sea una base ortonormal para $\mathbb{R}^n$, y entonces sea $U = [\mathbf{u}_1 \;\; \mathbf{u}_2 \;\; \cdots \;\; \mathbf{u}_n]$. Muestre que la primera columna de U^TAU es $\lambda_1\mathbf{e}_1$, donde $\mathbf{e}_1$ es la primera columna de la matriz identidad de $n \times n$.

b. El inciso (a) implica que U^TAU tiene la forma que se muestra a continuación. Explique por qué los valores propios de A_1 son $\lambda_2, \ldots, \lambda_n$. [*Sugerencia:* Vea los ejercicios suplementarios del capítulo 5.]

$$U^TAU = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{bmatrix}$$

[M] Cuando el miembro derecho de una ecuación $A\mathbf{x} = \mathbf{b}$ se cambia ligeramente —por ejemplo, a $A\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}$ para algún vector $\Delta\mathbf{b}$— la solución cambia de $\mathbf{x}$ a $\mathbf{x} + \Delta\mathbf{x}$, donde $\Delta\mathbf{x}$ satisface $A(\Delta\mathbf{x}) = \Delta\mathbf{b}$. El cociente $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$ se denomina **cambio relativo** de $\mathbf{b}$ (o **error relativo** en $\mathbf{b}$ cuando $\Delta\mathbf{b}$ representa el error posible en las entradas de $\mathbf{b}$). El cambio relativo en la solución es $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$. Cuando A es invertible, el **número de condición** de A, que se escribe cond(A), produce una cota para la magnitud del cambio relativo de $\mathbf{x}$:

---

[1] Si se permiten los números complejos, *toda* matriz A de $n \times n$ admite una factorización de Schur (compleja). $A = URU^{-1}$, donde R es triangular superior y U^{-1} es la transpuesta *conjugada* de U. Este hecho tan útil se analiza en *Matrix Analysis*, de Roger A. Horn y Charles R. Johnson (Cambridge: Cambridge University Press, 1985), págs. 79-100.

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \text{cond}(A)\cdot\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} \qquad (2)$$

En los ejercicios 17 a 20, resuelva $A\mathbf{x} = \mathbf{b}$ y $A(\Delta\mathbf{x}) = \Delta\mathbf{b}$, y muestre que (2) es válida en cada caso. (Vea el análisis de matrices *mal condicionadas* en los ejercicios 41, 42 y 43 de la sección 2.3.)

17. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 19.249 \\ 6.843 \end{bmatrix}$, $\Delta\mathbf{b} = \begin{bmatrix} .001 \\ -.003 \end{bmatrix}$

18. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} .500 \\ -1.407 \end{bmatrix}$, $\Delta\mathbf{b} = \begin{bmatrix} .001 \\ -.003 \end{bmatrix}$

19. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} .100 \\ 2.888 \\ -1.404 \\ 1.462 \end{bmatrix}$,

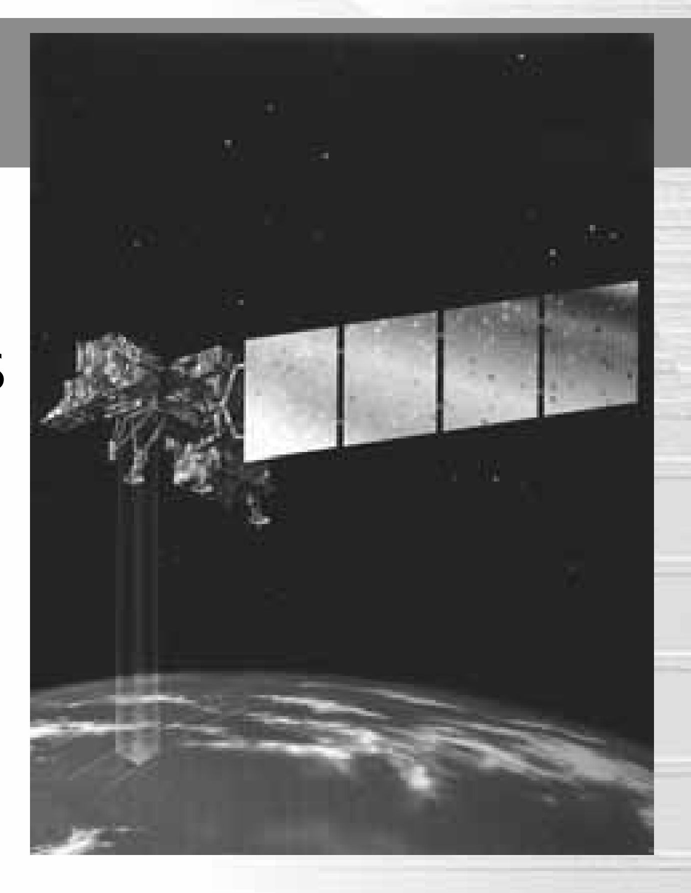
$\Delta\mathbf{b} = 10^{-4}\begin{bmatrix} .49 \\ -1.28 \\ 5.78 \\ 8.04 \end{bmatrix}$

20. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 4.230 \\ -11.043 \\ 49.991 \\ 69.536 \end{bmatrix}$,

$\Delta\mathbf{b} = 10^{-4}\begin{bmatrix} .27 \\ 7.76 \\ -3.77 \\ 3.93 \end{bmatrix}$

7

Matrices simétricas y formas cuadráticas



EJEMPLO INTRODUCTORIO

Procesamiento de imágenes multicanal

Dando la vuelta al mundo en poco más de 80 *minutos*, los dos satélites Landsat cruzan el cielo como un rayo silencioso con órbitas casi polares, graban imágenes del terreno y de las líneas costeras en franjas de 185 kilómetros de ancho. En periodos de 16 días, estos satélites pasan sobre casi todos los kilómetros cuadrados de la superficie terrestre, de modo que cualquier lugar se puede monitorear cada 8 días.

Las imágenes Landsat son útiles para muchos propósitos. Los desarrolladores y planificadores urbanos las usan para estudiar el ritmo y la dirección del crecimiento urbano, el desarrollo industrial, y otros cambios en el uso del suelo. Las comunidades rurales pueden analizar la humedad del suelo, clasificar la vegetación de áreas remotas, y localizar depósitos y corrientes de agua tierra adentro. Los gobiernos pueden detectar y estimar los daños debidos a desastres naturales, como incendios forestales, flujos de lava, inundaciones y huracanes. Las agencias de protección del medio ambiente pueden identificar la contaminación por emisiones de chimeneas y medir la temperatura del agua de lagos y ríos cercanos a plantas de energía.

Los sensores colocados a bordo de los satélites obtienen siete imágenes simultáneas de cualquier región de la Tierra que se vaya a estudiar. Estos sensores registran la energía en diferentes bandas de longitud de onda —tres en el espectro de luz visible y cuatro en las bandas de infrarrojo y térmico—. Cada imagen se digitaliza y archiva como una formación rectangular de números, donde cada número indica la intensidad de la señal en un pequeño punto (o *píxel*) correspondiente de la imagen. Cada una de las siete imágenes es un canal de una *imagen multicanal* o *multiespectral*.

Las siete imágenes Landsat de una región fija suelen contener mucha información redundante, puesto que algunas características aparecen en varias imágenes. Sin embargo, otras características, por su color o temperatura, pueden reflejar luz que registran únicamente uno o dos sensores. Una meta del procesamiento de imágenes multicanal es la de visualizar los datos de manera que la información se extraiga de mejor modo que estudiando cada imagen por separado.

El *análisis de componentes principales* es una manera efectiva de eliminar información redundante y de

proporcionar en una sola o en dos imágenes compuestas la mayor parte de la información proveniente de los datos iniciales. A grandes rasgos, el objetivo principal es encontrar una combinación lineal especial de las imágenes, es decir, una lista de pesos que combinen en cada píxel los siete valores correspondientes de las imágenes en un nuevo valor. Los pesos se eligen de tal manera que hagan al intervalo de intensidades de luz —la *varianza de la escena*— de la imagen compuesta (llamada *primera componente principal*) mayor que en cualquiera de las imágenes originales. También se pueden estructurar imágenes adicionales de *componentes*, aplicando criterios que se explicarán en la sección 7.5.

El análisis de componentes principales se ilustra en las siguientes fotografías, tomadas sobre el valle Railroad en Nevada, EUA. En (a), (b) y (c) se muestran las imágenes provenientes de tres bandas espectrales Landsat. La información total de las tres bandas se reacomoda en tres imágenes de componentes (d), (e) y (f). El primer componente (d) despliega (o "explica") 93.5% de la varianza de la escena presente en los datos iniciales. De esta manera, los datos iniciales de tres canales se han reducido a datos de un canal, con una pérdida en algún sentido de sólo el 6.5% de la varianza de la escena.

La empresa Earth Satellite Corporation de Rockville, Maryland, que amablemente proporcionó las fotografías mostradas, está experimentando con imágenes de 224 bandas espectrales individuales. El análisis de componentes principales, que resulta indispensable al tratar con tales conjuntos masivos de datos, a menudo reduce los datos a aproximadamente 15 componentes principales utilizables.

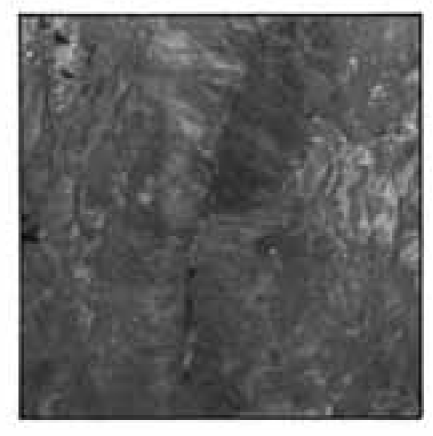
(a) Banda espectral 1: Azul visible.

(b) Banda espectral 4: Casi infrarrojo.

(c) Banda espectral 7: Infrarrojo medio.

(d) Componente principal 1: 93.5%.

(e) Componente principal 2: 5.3%.



(f) Componente principal 3: 1.2%.

Las matrices simétricas surgen en las aplicaciones, de una u otra manera, con mayor frecuencia que cualquier otra clase importante de matrices. La teoría es hermosa y rica, y depende, esencialmente, tanto de la técnica de diagonalización del capítulo 5 como de la ortogonalidad del capítulo 6. La diagonalización de una matriz simétrica, descrita en la sección 7.1, es el fundamento para el análisis presentado en las secciones 7.2 y 7.3 relativas a las formas cuadráticas. La sección 7.3, a su vez, es necesaria para comprender las dos secciones finales que tratan acerca de la descomposición en valores singulares y sobre el procesamiento de imágenes descrito en el ejemplo introductorio. A lo largo de este capítulo, los vectores y matrices tienen entradas reales.

7.1 DIAGONALIZACIÓN DE MATRICES SIMÉTRICAS

Una matriz **simétrica** es una matriz A tal que $A^T = A$. Una matriz de este tipo es necesariamente cuadrada. Sus entradas en la diagonal principal son arbitrarias, pero sus otras entradas ocurren en pares —en lados opuestos de la diagonal principal.

EJEMPLO 1 De las siguientes matrices, únicamente las tres primeras son simétricas:

$$\text{Simétricas:} \quad \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{bmatrix}, \quad \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$$

$$\text{No simétricas:} \quad \begin{bmatrix} 1 & -3 \\ 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -4 & 0 \\ -6 & 1 & -4 \\ 0 & -6 & 1 \end{bmatrix}, \quad \begin{bmatrix} 5 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$$

Para comenzar el estudio de las matrices simétricas, es útil repasar el proceso de diagonalización visto en la sección 5.3.

EJEMPLO 2 De ser posible, diagonalice la matriz $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

Solución La ecuación característica de A es

$$0 = -\lambda^3 + 17\lambda^2 - 90\lambda + 144 = -(\lambda - 8)(\lambda - 6)(\lambda - 3)$$

Los cálculos estándar producen una base para cada espacio propio;

$$\lambda = 8\colon\ \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = 6\colon\ \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}; \qquad \lambda = 3\colon\ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Estos tres vectores conforman una base para $\mathbb{R}^3$, y pueden usarse como columnas para una matriz P que diagonalice A. Sin embargo, puede advertirse fácilmente que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un *conjunto ortogonal*, y P resultará más útil si sus columnas son ortonormales. Dado que un múltiplo diferente de cero de un vector propio sigue siendo un vector propio, es posible normalizar a $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ para producir los vectores propios unitarios.

$$\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Sean

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}, \quad D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Entonces $A = PDP^{-1}$, como de costumbre. Pero esta vez, dado que P es cuadrada y tiene columnas ortonormales, P es una matriz *ortogonal*, y P^{-1} es simplemente P^T. (Vea la sección 6.2.)

El teorema 1 explica por qué los vectores propios del ejemplo 2 son ortogonales —corresponden a valores propios distintos.

TEOREMA 1 Si A es simétrica, entonces cualesquiera dos vectores propios de espacios propios diferentes son ortogonales.

DEMOSTRACIÓN Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ vectores propios correspondientes a distintos valores propios, por ejemplo, λ_1 y λ_2. Para demostrar que $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, calcule

$$\begin{aligned} \lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 &= (\lambda_1 \mathbf{v}_1)^T \mathbf{v}_2 = (A\mathbf{v}_1)^T \mathbf{v}_2 && \text{Puesto que } \mathbf{v}_1 \text{ es un vector propio} \\ &= (\mathbf{v}_1^T A^T)\mathbf{v}_2 = \mathbf{v}_1^T (A\mathbf{v}_2) && \text{Puesto que } A^T = A \\ &= \mathbf{v}_1^T (\lambda_2 \mathbf{v}_2) && \text{Puesto que } \mathbf{v}_2 \text{ es un vector propio} \\ &= \lambda_2 \mathbf{v}_1^T \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2 \end{aligned}$$

Por lo que $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Pero $\lambda_1 - \lambda_2 \neq 0$, así $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. ■

El tipo especial de diagonalización del ejemplo 2 es crucial para la teoría de matrices simétricas. Se afirma que una matriz A es **diagonalizable ortogonalmente** si existe una matriz ortogonal P (con $P^{-1} = P^T$) y una matriz diagonal D tales que

$$A = PDP^T = PDP^{-1} \tag{1}$$

Para diagonalizar ortogonalmente una matriz de $n \times n$, deben encontrarse n vectores propios linealmente independientes y ortonormales. ¿Cuándo es posible esto? Si A es diagonalizable ortogonalmente como en (1), entonces

$$A^T = (PDP^T)^T = P^{TT} D^T P^T = PDP^T = A$$

Por lo tanto, A es simétrica. El teorema 2 muestra que, recíprocamente, toda matriz simétrica es diagonalizable ortogonalmente. La demostración es mucho más difícil y se omite aquí; la idea principal para efectuar una demostración se proporcionará después del teorema 3.

TEOREMA 2 Una matriz A de $n \times n$ es diagonalizable ortogonalmente si, y sólo si, A es una matriz simétrica.

Este teorema es bastante sorprendente, porque la experiencia del capítulo 5 sugeriría que, por lo general, es imposible saber cuándo una matriz es diagonalizable. Pero éste no es el caso con matrices simétricas.

El siguiente ejemplo trata una matriz cuyos valores propios no son todos diferentes.

EJEMPLO 3 Diagonalice ortogonalmente la matriz $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, cuya ecuación característica es

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2)$$

Solución Los cálculos usuales producen bases para los espacios propios:

$$\lambda = 7\colon \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = -2\colon \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

Aunque $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes, no son ortogonales. Recuerde, de la sección 6.2, que la proyección de $\mathbf{v}_2$ sobre $\mathbf{v}_1$ es $\frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$, y la componente de $\mathbf{v}_2$ ortogonal a $\mathbf{v}_1$ es

$$\mathbf{z}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{-1/2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

Entonces $\{\mathbf{v}_1, \mathbf{z}_2\}$ es un conjunto ortogonal en el espacio propio para $\lambda = 7$. (Observe que $\mathbf{z}_2$ es una combinación lineal de los vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$, así que $\mathbf{z}_2$ está en el espacio propio. Esta estructuración de $\mathbf{z}_2$ es precisamente el proceso Gram-Schmidt de la sección 6.4.) Puesto que el espacio propio es bidimensional (con bases $\mathbf{v}_1$ y $\mathbf{v}_2$), El conjunto ortogonal $\{\mathbf{v}_1, \mathbf{z}_2\}$ es una *base ortogonal* para el espacio propio, de acuerdo con el teorema de la base. (Vea la sección 2.9 o la 4.5.)

Al normalizar $\mathbf{v}_1$ y $\mathbf{z}_2$ se obtiene la siguiente base ortonormal para el espacio propio con $\lambda = 7$:

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$$

Una base ortonormal para el espacio propio con $\lambda = -2$ es

$$\mathbf{u}_3 = \frac{1}{\|2\mathbf{v}_3\|}2\mathbf{v}_3 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

De acuerdo con el teorema 1, $\mathbf{u}_3$ es ortogonal a los otros vectores propios $\mathbf{u}_1$ y $\mathbf{u}_2$. Por lo tanto $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortonormal. Sean

$$P = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Entonces P diagonaliza ortogonalmente a A, y $A = PDP^{-1}$.

En el ejemplo 3, el valor propio 7 tiene multiplicidad dos y el espacio propio es bidimensional. Este hecho no es accidental, como lo muestra el teorema siguiente.

El teorema espectral

El conjunto de valores propios de una matriz A se denomina ocasionalmente como *espectro* de A, y la siguiente descripción de los valores propios es llamada *teorema espectral*.

TEOREMA 3 **El teorema espectral para matrices simétricas**

Una matriz simétrica A de $n \times n$ tiene las siguientes propiedades:

a. A tiene n valores propios reales, contando multiplicidades.
b. La dimensión del espacio propio para cada valor propio λ es igual a la multiplicidad de λ como raíz de la ecuación característica.
c. Los espacios propios son mutuamente ortogonales, en el sentido de que los vectores propios correspondientes a valores propios diferentes son ortogonales.
d. A es diagonalizable ortogonalmente.

El inciso (a) se deriva del ejercicio 24 presentado en la sección 5.5. El inciso (b) se deduce fácilmente del inciso (d). (Vea el ejercicio 31.) El inciso (c) es el teorema 1. A causa de (a), puede darse una demostración de (d) usando el ejercicio 32 y la factorización de Schur analizada en el ejercicio suplementario 16 del capítulo 6. Se omiten los detalles.

Descomposición espectral

Suponga que $A = PDP^{-1}$, donde las columnas de P son vectores propios ortonormales $\mathbf{u}_1, \ldots, \mathbf{u}_n$ de A y los valores propios correspondientes $\lambda_1, \ldots, \lambda_n$ están en la matriz diagonal D. Entonces, como $P^{-1} = P^T$,

$$A = PDP^T = [\,\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_n\,] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

$$= [\,\lambda_1\mathbf{u}_1 \quad \cdots \quad \lambda_n\mathbf{u}_n\,] \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

A partir del desarrollo de columna-fila de un producto (teorema 10 de la sección 2.4), puede escribirse

$$A = \lambda_1 \mathbf{u}_1\mathbf{u}_1^T + \lambda_2 \mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n \mathbf{u}_n\mathbf{u}_n^T \qquad (2)$$

Esta representación de A se llama **descomposición espectral** de A porque divide a A en fragmentos determinados por el espectro (valores propios) de A. Cada término de (2) es una matriz de $n \times n$ de rango 1. Por ejemplo, cada columna de $\lambda_1\mathbf{u}_1\mathbf{u}_1^T$ es un múltiplo de $\mathbf{u}_1$. Más aún, cada matriz $\mathbf{u}_j\mathbf{u}_j^T$ es una **matriz de proyección** en el sentido de que para cada $\mathbf{x}$ en $\mathbb{R}^n$, el vector $(\mathbf{u}_j\mathbf{u}_j^T)\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre el subespacio generado por $\mathbf{u}_j$. (Vea el ejercicio 35.)

EJEMPLO 4 Estructure una descomposición espectral de la matriz A que tiene la diagonalización ortogonal

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}\begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Solución Denote las columnas de P mediante $\mathbf{u}_1$ y $\mathbf{u}_2$. Entonces

$$A = 8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T$$

Para verificar esta descomposición de A, calcule

$$\mathbf{u}_1\mathbf{u}_1^T = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}$$

$$\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}\begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix}$$

y

$$8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} 32/5 & 16/5 \\ 16/5 & 8/5 \end{bmatrix} + \begin{bmatrix} 3/5 & -6/5 \\ -6/5 & 12/5 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = A$$

NOTA NUMÉRICA

Cuando A es simétrica y no demasiado grande, los algoritmos de computadora modernos de alto rendimiento calculan con gran precisión vectores y valores propios. Estos algoritmos aplican a A una sucesión de transformaciones de semejanza en las que intervienen matrices ortogonales. Las entradas diagonales de las matrices transformadas convergen rápidamente hacia los valores propios de A. (Vea las notas numéricas de la sección 5.2.) Por lo general, el uso de matrices ortogonales evita que los errores numéricos se acumulen durante el proceso. Cuando A es simétrica, la sucesión de matrices ortogonales se combina para formar una matriz ortogonal cuyas columnas son vectores propios de A.

Una matriz no simétrica no puede tener un conjunto completo de vectores propios ortogonales, pero el algoritmo aún produce valores propios bastante precisos. Después de eso, se necesitan técnicas no ortogonales para calcular los vectores propios.

PROBLEMAS DE PRÁCTICA

1. Muestre que si A es una matriz simétrica, entonces A^2 es simétrica.
2. Muestre que si A es diagonalizable ortogonalmente, entonces A^2 también lo es.

7.1 EJERCICIOS

Determine cuáles de las matrices presentadas en los ejercicios 1 a 6 son simétricas.

1. $\begin{bmatrix} 3 & 5 \\ 5 & -7 \end{bmatrix}$

2. $\begin{bmatrix} -3 & 5 \\ -5 & 3 \end{bmatrix}$

3. $\begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix}$

4. $\begin{bmatrix} 0 & 8 & 3 \\ 8 & 0 & -2 \\ 3 & -2 & 0 \end{bmatrix}$

5. $\begin{bmatrix} -6 & 2 & 0 \\ 0 & -6 & 2 \\ 0 & 0 & -6 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \end{bmatrix}$

Determine cuáles de las matrices presentadas en los ejercicios 7 a 12 son ortogonales. Si son ortogonales, encuentre el inverso.

7. $\begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}$

8. $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

9. $\begin{bmatrix} -5 & 2 \\ 2 & 5 \end{bmatrix}$

10. $\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$

11. $\begin{bmatrix} 2/3 & 2/3 & 1/3 \\ 0 & 1/\sqrt{5} & -2/\sqrt{5} \\ \sqrt{5}/3 & -4/\sqrt{45} & -2/\sqrt{45} \end{bmatrix}$

12. $\begin{bmatrix} .5 & .5 & -.5 & -.5 \\ -.5 & .5 & -.5 & .5 \\ .5 & .5 & .5 & .5 \\ -.5 & .5 & .5 & -.5 \end{bmatrix}$

Diagonalice ortogonalmente las matrices de los ejercicios 13 a 22, proporcione una matriz ortogonal P y una matriz diagonal D. Para ahorrarle tiempo, los valores propios de los ejercicios 17 a 24 son: (17) 5, 2, −2; (18) 25, 3, −50; (19) 7, −2; (20) 13, 7, 1; (21) 9, 5, 1; (22) 2, 0.

13. $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 16 & -4 \\ -4 & 1 \end{bmatrix}$

16. $\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$

18. $\begin{bmatrix} -2 & -36 & 0 \\ -36 & -23 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

19. $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

20. $\begin{bmatrix} 7 & -4 & 4 \\ -4 & 5 & 0 \\ 4 & 0 & 9 \end{bmatrix}$

21. $\begin{bmatrix} 4 & 1 & 3 & 1 \\ 1 & 4 & 1 & 3 \\ 3 & 1 & 4 & 1 \\ 1 & 3 & 1 & 4 \end{bmatrix}$

22. $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

23. Sean $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Verifique si 2 es un valor propio de A y $\mathbf{v}$ un vector propio. Luego diagonalice ortogonalmente a A.

24. Sean $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$, y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. Compruebe que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios de A. Después, diagonalice ortogonalmente a A.

En los ejercicios 25 y 26, señale cada enunciado como verdadero o falso. Justifique sus respuestas.

25. a. Una matriz de $n \times n$ que es diagonalizable ortogonalmente debe ser simétrica.

b. Si $A^T = A$, y si los vectores $\mathbf{u}$ y $\mathbf{v}$ satisfacen $A\mathbf{u} = 3\mathbf{u}$ y $A\mathbf{v} = 4\mathbf{v}$, entonces $\mathbf{u} \cdot \mathbf{v} = 0$.

c. Una matriz simétrica de $n \times n$ tiene n valores propios reales distintos.

d. Para un $\mathbf{v}$ diferente de cero en $\mathbb{R}^n$, la matriz $\mathbf{v}\mathbf{v}^T$ se denomina matriz de proyección.

26. a. Toda matriz simétrica es diagonalizable ortogonalmente.

b. Si $B = PDP^T$, donde $P^T = P^{-1}$ y D es una matriz diagonal, entonces B es una matriz simétrica.

c. Una matriz ortogonal es diagonalizable ortogonalmente.

d. La dimensión de un espacio propio de una matriz simétrica equivale a la multiplicidad del valor propio correspondiente.

27. Suponga que A es una matriz simétrica de $n \times n$ y que B es cualquier matriz de $n \times m$. Muestre que B^TAB, B^TB, y BB^T son matrices simétricas.

28. Muestre que si A es una matriz simétrica de $n \times n$, entonces $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$ para todas $\mathbf{x}$, $\mathbf{y}$ en $\mathbb{R}^n$.

29. Suponga que A es invertible y diagonalizable ortogonalmente. Explique por qué A^{-1} también es diagonalizable ortogonalmente.

30. Suponga que tanto A como B son diagonalizables ortogonalmente y que $AB = BA$. Explique por qué AB también es diagonalizable ortogonalmente.

31. Sea $A = PDP^{-1}$, donde P es ortogonal y D es diagonal, y sea λ un valor propio de A con multiplicidad k. Entonces λ aparece k veces en la diagonal de D. Explique por qué la dimensión del espacio propio para λ es k.

32. Suponga que $A = PRP^{-1}$, donde P es ortogonal y R es triangular superior. Muestre que si A es simétrica, entonces R es simétrica y, por lo tanto, es realmente una matriz diagonal.

33. Construya una descomposición espectral de la A del ejemplo 2.

34. Construya una descomposición espectral de la A del ejemplo 3.

35. Sea $\mathbf{u}$ un vector unitario en $\mathbb{R}^n$, y sea $B = \mathbf{u}\mathbf{u}^T$.

a. Dado cualquier $\mathbf{x}$ en $\mathbb{R}^n$, calcule $B\mathbf{x}$ y muestre que $B\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre $\mathbf{u}$, como se describió en la sección 6.2.

b. Muestre que B es una matriz simétrica y que $B^2 = B$.

c. Muestre que $\mathbf{u}$ es un vector propio de B. ¿Cuál es el valor propio correspondiente?

36. Sea B una matriz simétrica de $n \times n$ tal que $B^2 = B$. Cualquier matriz de este tipo se denomina **matriz de proyección** (o **matriz de proyección ortogonal**). Dado cualquier $\mathbf{y}$ en $\mathbb{R}^n$, sean $\hat{\mathbf{y}} = B\mathbf{y}$ y $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

a. Muestre que $\mathbf{z}$ es ortogonal a $\hat{\mathbf{y}}$.

b. Sea W el espacio de columnas de B. Muestre que $\mathbf{y}$ es la suma de un vector en W y un vector en $W^{\perp}$. ¿Por qué demuestra esto que $B\mathbf{y}$ es la proyección ortogonal de $\mathbf{y}$ sobre el espacio de columnas de B?

[M] Diagonalice ortogonalmente las matrices de los ejercicios 37 a 40. Para practicar los métodos de esta sección, no use una rutina de vectores propios del programa de matrices. En vez de eso, utilice el programa para encontrar los valores propios, y, para cada valor propio λ, encuentre una base ortonormal para Nul$(A - \lambda I)$, como en los ejemplos 2 y 3.

37. $\begin{bmatrix} 5 & 2 & 9 & -6 \\ 2 & 5 & -6 & 9 \\ 9 & -6 & 5 & 2 \\ -6 & 9 & 2 & 5 \end{bmatrix}$

38. $\begin{bmatrix} .38 & -.18 & -.06 & -.04 \\ -.18 & .59 & -.04 & .12 \\ -.06 & -.04 & .47 & -.12 \\ -.04 & .12 & -.12 & .41 \end{bmatrix}$

39. $\begin{bmatrix} .31 & .58 & .08 & .44 \\ .58 & -.56 & .44 & -.58 \\ .08 & .44 & .19 & -.08 \\ .44 & -.58 & -.08 & .31 \end{bmatrix}$

40. $\begin{bmatrix} 10 & 2 & 2 & -6 & 9 \\ 2 & 10 & 2 & -6 & 9 \\ 2 & 2 & 10 & -6 & 9 \\ -6 & -6 & -6 & 26 & 9 \\ 9 & 9 & 9 & 9 & -19 \end{bmatrix}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. $(A^2)^T = (AA)^T = A^TA^T$, de acuerdo con una propiedad de las transpuestas. Por hipótesis, $A^T = A$. De modo que $(A^2)^T = AA = A^2$, lo cual muestra que A^2 es simétrica.

2. Si A es diagonalizable ortogonalmente, entonces A es simétrica, según el teorema 2. De acuerdo con el problema de práctica 1, A^2 es simétrica y, por lo tanto, es diagonalizable ortogonalmente (teorema 2).

7.2 FORMAS CUADRÁTICAS

Hasta ahora, la atención en este texto se ha enfocado en ecuaciones lineales, con excepción de las sumas de cuadrados encontradas en el capítulo 6 al calcular $\mathbf{x}^T\mathbf{x}$. Tales sumas y otras expresiones más generales, llamadas *formas cuadráticas*, se presentan a menudo en aplicaciones de álgebra lineal a la ingeniería (en criterios de diseño y optimización) y al procesamiento de señales (como potencia de ruido de salida). También surgen, por

ejemplo, en física (como energías potencial y cinética), en geometría diferencial (como la curvatura normal de las superficies), en economía (como funciones de utilidad), y en estadística (en elipsoides de confianza). Algunos de los antecedentes matemáticos para encarar tales aplicaciones fluyen con facilidad a partir del trabajo realizado en este texto con las matrices simétricas.

Una **forma cuadrática** en $\mathbb{R}^n$ es una función Q definida en $\mathbb{R}^n$ cuyo valor en un vector $\mathbf{x}$ en $\mathbb{R}^n$ puede calcularse mediante una expresión de la forma $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, donde A es una matriz simétrica de $n \times n$. La matriz A se denomina **matriz de la forma cuadrática**.

El ejemplo más sencillo de una forma cuadrática diferente de cero es $Q(\mathbf{x}) = \mathbf{x}^TI\mathbf{x} = \|\mathbf{x}\|^2$. Los ejemplos 1 y 2 muestran la conexión que hay entre cualquier matriz simétrica A y la forma cuadrática $\mathbf{x}^TA\mathbf{x}$.

EJEMPLO 1 Sea $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Calcule $\mathbf{x}^TA\mathbf{x}$ para las siguientes matrices:

a. $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$ b. $A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

Solución

a. $\mathbf{x}^TA\mathbf{x} = [\, x_1 \quad x_2 \,]\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\, x_1 \quad x_2 \,]\begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2.$

b. Existen dos entradas -2 en A. Observe cómo aparecen en los cálculos. La entrada (1, 2) de A está en negritas.

$$\begin{aligned} \mathbf{x}^TA\mathbf{x} &= [\, x_1 \quad x_2 \,]\begin{bmatrix} 3 & \mathbf{-2} \\ -2 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\, x_1 \quad x_2 \,]\begin{bmatrix} 3x_1 - \mathbf{2}x_2 \\ -2x_1 + 7x_2 \end{bmatrix} \\ &= x_1(3x_1 - \mathbf{2}x_2) + x_2(-2x_1 + 7x_2) \\ &= 3x_1^2 - \mathbf{2}x_1x_2 - 2x_2x_1 + 7x_2^2 \\ &= 3x_1^2 - 4x_1x_2 + 7x_2^2 \end{aligned}$$

La presencia de $-4x_1x_2$ en la forma cuadrática del ejemplo 1(b) se debe a las entradas -2 fuera de la diagonal en la matriz A. En contraste, la forma cuadrática asociada con la matriz diagonal A del ejemplo 1(a) no tiene ningún término de *producto cruzado* x_1x_2.

EJEMPLO 2 Para $\mathbf{x}$ en $\mathbb{R}^3$, sea $Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$. Escriba esta forma cuadrática como $\mathbf{x}^TA\mathbf{x}$.

Solución Los coeficientes de x_1^2, x_2^2, x_3^2 van en la diagonal de A. Para hacer simétrica a A, el coeficiente de x_ix_j para $i \neq j$ debe dividirse uniformemente entre las (i, j)-ésimas y (j, i)-ésimas entradas de A. El coeficiente de x_1x_3 es cero. Se comprueba fácilmente que

$$Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x} = [\, x_1 \quad x_2 \quad x_3 \,]\begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

EJEMPLO 3 Sea $Q(\mathbf{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$. Calcule el valor de $Q(\mathbf{x})$ para $\mathbf{x} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -2 \end{bmatrix}$, y $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

Solución

$$Q(-3, 1) = (-3)^2 - 8(-3)(1) - 5(1)^2 = 28$$
$$Q(2, -2) = (2)^2 - 8(2)(-2) - 5(-2)^2 = 16$$
$$Q(1, -3) = (1)^2 - 8(1)(-3) - 5(-3)^2 = -20$$

En algunos casos es más fácil usar formas cuadráticas cuando no tienen términos de producto cruzado, esto es, cuando la matriz de la forma cuadrática es una matriz diagonal. Por fortuna, el término de producto cruzado puede eliminarse mediante un cambio de variable adecuado.

Cambio de variable en una forma cuadrática

Si $\mathbf{x}$ representa un vector variable en $\mathbb{R}^n$, entonces un **cambio de variable** es una ecuación de la forma

$$\mathbf{x} = P\mathbf{y}, \qquad \text{o de manera equivalente,} \qquad \mathbf{y} = P^{-1}\mathbf{x} \tag{1}$$

donde P es una matriz invertible e $\mathbf{y}$ es un nuevo vector variable en $\mathbb{R}^n$. Aquí $\mathbf{y}$ es el vector de coordenadas de $\mathbf{x}$ relativo a la base de $\mathbb{R}^n$ determinada por las columnas de P. (Vea la sección 4.4.)

Si se aplica el cambio de variable (1) en una forma cuadrática $\mathbf{x}^TA\mathbf{x}$, entonces

$$\mathbf{x}^TA\mathbf{x} = (P\mathbf{y})^TA(P\mathbf{y}) = \mathbf{y}^TP^TAP\mathbf{y} = \mathbf{y}^T(P^TAP)\mathbf{y} \tag{2}$$

y la nueva matriz de la forma cuadrática es P^TAP. Si P diagonaliza *ortogonalmente* a A, entonces $P^T = P^{-1}$ y $P^TAP = P^{-1}AP = D$. ¡La matriz de la nueva forma cuadrática es diagonal! Ésta es la estrategia del siguiente ejemplo.

EJEMPLO 4 Efectúe un cambio de variable que transforme la forma cuadrática del ejemplo 3 en una forma cuadrática sin términos de producto cruzado.

Solución La matriz de la forma cuadrática del ejemplo 3 es

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

El primer paso consiste en diagonalizar ortogonalmente a A. Sus valores propios resultan ser $\lambda = 3$ y $\lambda = -7$. Los vectores propios unitarios asociados son

$$\lambda = 3\text{:} \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \qquad \lambda = -7\text{:} \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

Estos vectores son automáticamente ortogonales (porque corresponden a valores propios distintos) y, por lo tanto, proporcionan una base ortonormal para $\mathbb{R}^2$. Sean

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

Entonces $A = PDP^{-1}$ y $D = P^{-1}AP = P^TAP$, como fue señalado antes. Un cambio de variable apropiado es

$$\mathbf{x} = P\mathbf{y}, \qquad \text{donde } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Entonces

$$\begin{aligned} x_1^2 - 8x_1x_2 - 5x_2^2 &= \mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) \\ &= \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} \\ &= 3y_1^2 - 7y_2^2 \end{aligned}$$

Para ilustrar el significado de la igualdad de formas cuadráticas dado en el ejemplo 4, se puede calcular $Q(\mathbf{x})$ para $\mathbf{x} = (2, -2)$ usando la nueva forma cuadrática. Primero, como $\mathbf{x} = P\mathbf{y}$, se tiene que

$$\mathbf{y} = P^{-1}\mathbf{x} = P^T\mathbf{x}$$

así que

$$\mathbf{y} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

Por lo tanto,

$$\begin{aligned} 3y_1^2 - 7y_2^2 &= 3(6/\sqrt{5})^2 - 7(-2/\sqrt{5})^2 = 3(36/5) - 7(4/5) \\ &= 80/5 = 16 \end{aligned}$$

Éste es el valor de $Q(\mathbf{x})$ en el ejemplo 3 cuando $\mathbf{x} = (2, -2)$. Vea la figura 1.

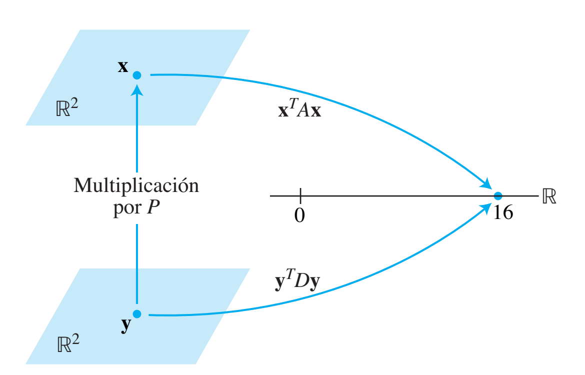


FIGURA 1 Cambio de variable en $\mathbf{x}^T A\mathbf{x}$.

En el ejemplo 4 se ilustra el teorema siguiente. La demostración del teorema se dio, en lo esencial, antes del ejemplo 4.

TEOREMA 4 **El teorema de los ejes principales**

Sea A una matriz simétrica de $n \times n$. Entonces existe un cambio ortogonal de variable, $\mathbf{x} = P\mathbf{y}$, que transforma la forma cuadrática $\mathbf{x}^T A\mathbf{x}$ en una forma cuadrática $\mathbf{y}^T D\mathbf{y}$ sin términos de producto cruzado.

Las columnas de la P del teorema se llaman **ejes principales** de la forma cuadrática $\mathbf{x}^T A\mathbf{x}$. El vector $\mathbf{y}$ es el vector de coordenadas de $\mathbf{x}$ relativo a la base ortonormal de $\mathbb{R}^n$ dada por estos ejes principales.

Una perspectiva geométrica de los ejes principales

Suponga que $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, donde A es una matriz simétrica invertible de 2×2, y sea c una constante. Puede mostrarse que el conjunto de todas las $\mathbf{x}$ en $\mathbb{R}^2$ que satisface

$$\mathbf{x}^T A\mathbf{x} = c \tag{3}$$

o corresponde a una elipse (o círculo), a una hipérbola, a dos líneas que se intersecan, o a un solo punto, o no contiene ningún punto. Si A es una matriz diagonal, la gráfica está en *posición estándar*, como en la figura 2. Si A es una matriz no diagonal, la gráfica de (3) está girada hasta salirse de la posición estándar, como en la figura 3 (pág. 460). Encontrar los *ejes principales* (determinados por los vectores propios de A) equivale a encontrar un nuevo sistema de coordenadas con respecto al cual la gráfica está en posición estándar.

La hipérbola de la figura 3(b) es la gráfica de la ecuación $\mathbf{x}^T A\mathbf{x} = 16$, donde A es la matriz del ejemplo 4. El eje y_1 positivo de la figura 3(b) está en la dirección de la primera columna de la P del ejemplo 4, y el eje y_2 positivo está en la dirección de la segunda columna de P.

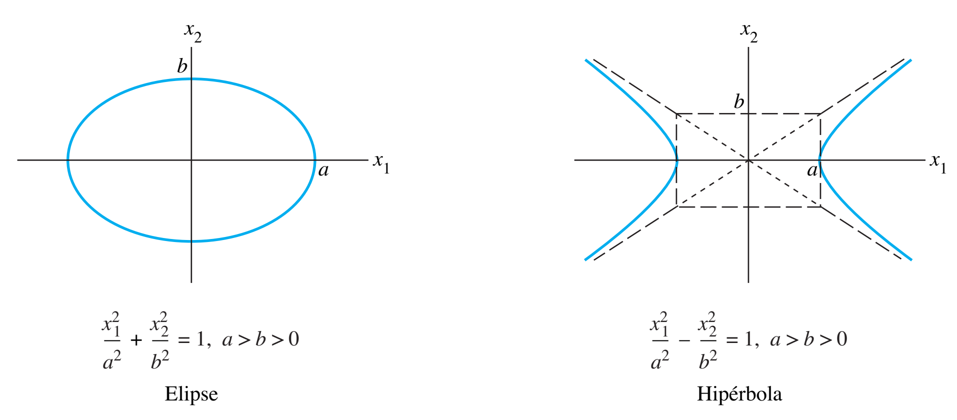
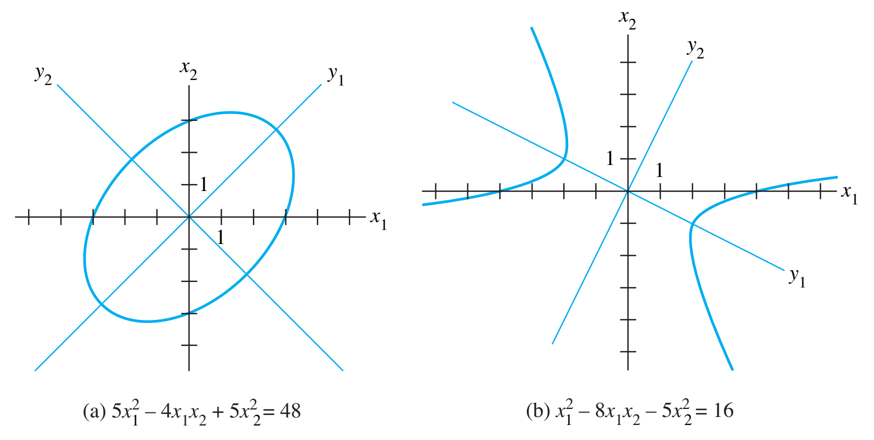


FIGURA 2 Una elipse y una hipérbola en posición estándar.

EJEMPLO 5 La elipse de la figura 3(a) es la gráfica de la ecuación $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$. Encuentre un cambio de variable que elimine de la ecuación el término del producto cruzado.

Solución La matriz de la forma cuadrática es $A = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$. Los valores propios de A resultan ser 3 y 7, con vectores propios unitarios correspondientes

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$



(a) $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$ (b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

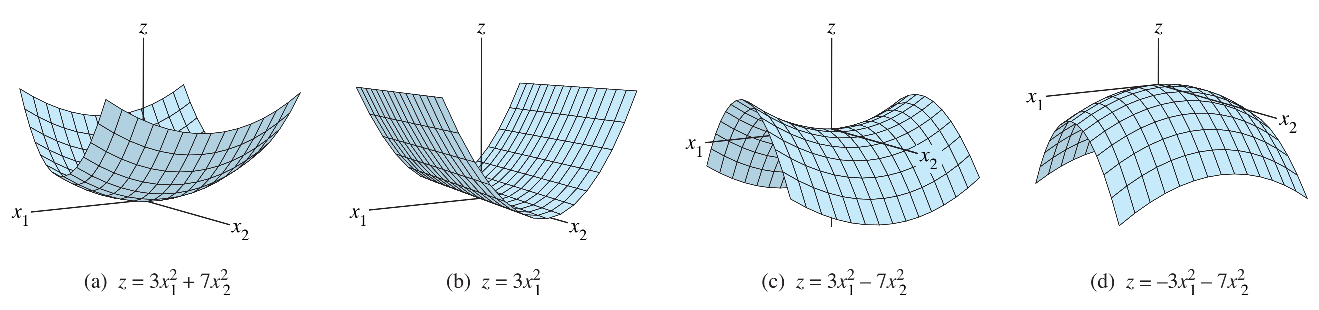
FIGURA 3 Una elipse y una hipérbola que *no están* en posición estándar.

Sea $P = [\,\mathbf{u}_1 \quad \mathbf{u}_2\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. Entonces P diagonaliza ortogonalmente a A, así que el cambio de variable $\mathbf{x} = P\mathbf{y}$ produce la forma cuadrática $\mathbf{y}^T D\mathbf{y} = 3y_1^2 + 7y_2^2$. Los nuevos ejes para este cambio de variable se muestran en la figura 3(a).

Clasificación de formas cuadráticas

Cuando A es una matriz de $n \times n$, la forma cuadrática $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ es una función de valores reales con dominio $\mathbb{R}^n$. Se distinguen varias clases importantes de formas cuadráticas por el tipo de *valores* que asumen para diversos $\mathbf{x}$.

En la figura 4 se muestran las gráficas de cuatro formas cuadráticas. Para cada punto $\mathbf{x} = (x_1, x_2)$ del dominio de una forma cuadrática Q, se traza un punto (x_1, x_2, z), donde $z = Q(\mathbf{x})$. Observe que excepto en $\mathbf{x} = \mathbf{0}$, todos los valores de $Q(\mathbf{x})$ son positivos en la figura 4(a) y negativos en la figura 4(d). Las secciones transversales horizontales de las gráficas son elipses en las figuras 4(a) y 4(d) e hipérbolas en 4(c).



(a) $z = 3x_1^2 + 7x_2^2$ (b) $z = 3x_1^2$ (c) $z = 3x_1^2 - 7x_2^2$ (d) $z = -3x_1^2 - 7x_2^2$

FIGURA 4 Gráficas de formas cuadráticas.

Los sencillos ejemplos 2×2 de la figura 4 ilustran las siguientes definiciones.

DEFINICIÓN

Una forma cuadrática Q es:

a. **definida positiva** si $Q(\mathbf{x}) > 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

b. **definida negativa** si $Q(\mathbf{x}) < 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

c. **indefinida** si $Q(\mathbf{x})$ toma valores tanto positivos como negativos.

Asimismo, se afirma que Q es **semidefinida positiva** si $Q(\mathbf{x}) \geq 0$ para toda $\mathbf{x}$, y Q es **semidefinida negativa** si $Q(\mathbf{x}) \leq 0$ para toda $\mathbf{x}$. Las formas cuadráticas de los incisos (a) y (b) de la figura 4 son ambas semidefinidas positivas.

El teorema 5 caracteriza algunas formas cuadráticas en términos de los valores propios.

TEOREMA 5

Formas cuadráticas y valores propios

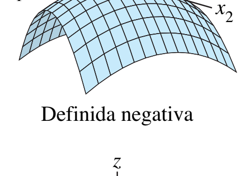
Sea A una matriz simétrica de $n \times n$. Entonces una forma cuadrática $\mathbf{x}^T A\mathbf{x}$ es:

a. definida positiva si, y sólo si, todos los valores propios de A son positivos,

b. definida negativa si, y sólo si, todos los valores propios de A son negativos, o

c. indefinida si, y sólo si, A tiene valores propios tanto positivos como negativos.

Definida positiva



Definida negativa

Indefinida

DEMOSTRACIÓN De acuerdo con el teorema de los ejes principales, existe un cambio de variable ortogonal $\mathbf{x} = P\mathbf{y}$ tal que

$$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \tag{4}$$

donde $\lambda_1, \ldots, \lambda_n$ son los valores propios de A. Como P es invertible, existe una correspondencia uno a uno entre todos los $\mathbf{x}$ diferentes de cero y todos los $\mathbf{y}$ distintos de cero. Entonces los valores de $Q(\mathbf{x})$ para $\mathbf{x} \neq \mathbf{0}$ coinciden con los valores de la expresión del lado derecho de (4), que están obviamente controlados *por* los signos de los valores propios $\lambda_1, \ldots, \lambda_n$, de las tres maneras descritas en el teorema.

EJEMPLO 6 ¿Es $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ definida positiva?

Solución Por todos los signos de suma, la forma "parece" definida positiva. Pero la matriz de la forma es

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

y los valores propios de A resultan ser 5, 2, y -1. Así que Q es una forma cuadrática indefinida, no definida positiva.

WEB

La clasificación de una forma cuadrática a menudo se propaga a la matriz de la forma. Entonces una **matriz definida positiva** A es una matriz *simétrica* para la cual la forma cuadrática $\mathbf{x}^T A\mathbf{x}$ es definida positiva. Los otros términos, como **matriz semidefinida positiva**, se definen de manera análoga.

NOTA NUMÉRICA

Un modo rápido de determinar si una matriz simétrica A es definida positiva es intentar factorizar A de la forma $A = R^T R$, donde R es triangular superior con entradas diagonales positivas. (Un algoritmo ligeramente modificado para una factorización LU es uno de los enfoques posibles.) Una *factorización Cholesky* de este tipo es posible si, y sólo si, A es definida positiva. Vea el ejercicio suplementario 7.

PROBLEMA DE PRÁCTICA

Describa una matriz A semidefinida positiva en términos de sus valores propios.

7.2 EJERCICIOS

1. Determine la forma cuadrática $\mathbf{x}^T A\mathbf{x}$, cuando $A = \begin{bmatrix} 5 & 1/3 \\ 1/3 & 1 \end{bmatrix}$ y

a. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ b. $\mathbf{x} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$ c. $\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

2. Determine la forma cuadrática $\mathbf{x}^T A\mathbf{x}$, para $A = \begin{bmatrix} 4 & 3 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ y

a. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ b. $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$ c. $\mathbf{x} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$

3. Encuentre la matriz de la forma cuadrática. Suponga que $\mathbf{x}$ está en $\mathbb{R}^2$.

a. $10x_1^2 - 6x_1x_2 - 3x_2^2$ b. $5x_1^2 + 3x_1x_2$

4. Encuentre la matriz de la forma cuadrática. Suponga que $\mathbf{x}$ está en $\mathbb{R}^2$.

a. $20x_1^2 + 15x_1x_2 - 10x_2^2$ b. x_1x_2

5. Encuentre la matriz de la forma cuadrática. Suponga que $\mathbf{x}$ está en $\mathbb{R}^3$.

a. $8x_1^2 + 7x_2^2 - 3x_3^2 - 6x_1x_2 + 4x_1x_3 - 2x_2x_3$

b. $4x_1x_2 + 6x_1x_3 - 8x_2x_3$

6. Encuentre la matriz de la forma cuadrática. Suponga que $\mathbf{x}$ está en $\mathbb{R}^3$.

a. $5x_1^2 - x_2^2 + 7x_3^2 + 5x_1x_2 - 3x_1x_3$

b. $x_3^2 - 4x_1x_2 + 4x_2x_3$

7. Realice un cambio de variable, $\mathbf{x} = P\mathbf{y}$, que transforme la forma cuadrática ${x_1}^2 + 10x_1x_2 + {x_2}^2$ en una forma cuadrática sin término de producto cruzado. Dé P y la nueva forma cuadrática.

8. Sea A la matriz de la forma cuadrática

$$9x_1^2 + 7x_2^2 + 11x_3^2 - 8x_1x_2 + 8x_1x_3$$

Puede mostrarse que los valores propios de A son 3, 9 y 15. Encuentre una matriz ortogonal P tal que el cambio de variable $\mathbf{x} = P\mathbf{y}$ transforme $\mathbf{x}^T A\mathbf{x}$ en una forma cuadrática sin término de producto cruzado. Dé P y la nueva forma cuadrática.

Clasifique las formas cuadráticas de los ejercicios 9 a 18. Después realice un cambio de variable, $\mathbf{x} = P\mathbf{y}$, que transforme la forma cuadrática en una forma cuadrática sin término de producto cruzado. Escriba la nueva forma cuadrática. Estructure P usando los métodos de la sección 7.1.

9. $3x_1^2 - 4x_1x_2 + 6x_2^2$

10. $9x_1^2 - 8x_1x_2 + 3x_2^2$

11. $2x_1^2 + 10x_1x_2 + 2x_2^2$

12. $-5x_1^2 + 4x_1x_2 - 2x_2^2$

13. $x_1^2 - 6x_1x_2 + 9x_2^2$

14. $8x_1^2 + 6x_1x_2$

15. **[M]** $-2x_1^2 - 6x_2^2 - 9x_3^2 - 9x_4^2 + 4x_1x_2 + 4x_1x_3 + 4x_1x_4 + 6x_3x_4$

16. **[M]** $4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_4^2 + 3x_1x_2 + 3x_3x_4 - 4x_1x_4 + 4x_2x_3$

17. **[M]** $x_1^2 + x_2^2 + x_3^2 + x_4^2 + 9x_1x_2 - 12x_1x_4 + 12x_2x_3 + 9x_3x_4$

18. **[M]** $11x_1^2 - x_2^2 - 12x_1x_2 - 12x_1x_3 - 12x_1x_4 - 2x_3x_4$

19. ¿Cuál es el mayor valor posible de la forma cuadrática $5x_1^2 + 8x_2^2$ si $\mathbf{x} = (x_1, x_2)$ y $\mathbf{x}^T\mathbf{x} = 1$, es decir, si $x_1^2 + x_2^2 = 1$? (Pruebe con algunos ejemplos de $\mathbf{x}$.)

20. ¿Cuál es el mayor valor de la forma cuadrática $5x_1^2 - 3x_2^2$ si $\mathbf{x}^T\mathbf{x} = 1$?

En los ejercicios 21 y 22, las matrices son de $n \times n$ y los vectores están en $\mathbb{R}^n$. Señale cada enunciado como verdadero o falso. Justifique sus respuestas.

21. a. La matriz de una forma cuadrática es una matriz simétrica.

b. Una forma cuadrática no tendrá términos de producto cruzado si, y sólo si, la matriz de la forma cuadrática es una matriz diagonal.

c. Los ejes principales de una forma cuadrática $\mathbf{x}^TA\mathbf{x}$ son vectores propios de A.

d. Una forma cuadrática Q definida positiva satisface $Q(\mathbf{x}) > 0$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

e. Si todos los valores propios de una matriz simétrica A son positivos, entonces la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ es definida positiva.

f. Una factorización Cholesky de una matriz simétrica A tiene la forma $A = R^TR$, para una matriz triangular superior R con entradas diagonales positivas.

22. a. La expresión $\|\mathbf{x}\|^2$ es una forma cuadrática.

b. Si A es simétrica y P es una matriz ortogonal, entonces el cambio de variable $\mathbf{x} = P\mathbf{y}$ transforma $\mathbf{x}^TA\mathbf{x}$ en una forma cuadrática sin términos de producto cruzado.

c. Si A es una matriz simétrica de 2×2, entonces el conjunto de $\mathbf{x}$ tales que $\mathbf{x}^TA\mathbf{x} = c$ (para una constante c) corresponde a un círculo, a una elipse o a una hipérbola.

d. Una forma cuadrática indefinida es una forma semidefinida positiva o una forma semidefinida negativa.

e. Si A es simétrica y la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ sólo tiene valores negativos para $\mathbf{x} \neq \mathbf{0}$, entonces todos los valores propios de A son negativos.

En los ejercicios 23 y 24 se muestra cómo clasificar una forma cuadrática $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, donde $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ y $\det A \neq 0$, sin encontrar los valores propios de A.

23. Si λ_1 y λ_2 son los valores propios de A, entonces el polinomio característico de A puede escribirse de dos maneras: $\det(A - \lambda I)$ y $(\lambda - \lambda_1)(\lambda - \lambda_2)$. Utilice este hecho para mostrar que $\lambda_1 + \lambda_2 = a + d$ (las entradas diagonales de A) y $\lambda_1\lambda_2 = \det A$.

24. Verifique las siguientes afirmaciones,

a. Q es definida positiva si $\det A > 0$ y $a > 0$.

b. Q es definida negativa si $\det A > 0$ y $a < 0$.

c. Q es indefinida si $\det A < 0$.

25. Muestre que si B es de $m \times n$, entonces B^TB es semidefinida positiva; y si B es de $n \times n$ e invertible, entonces B^TB es definida positiva.

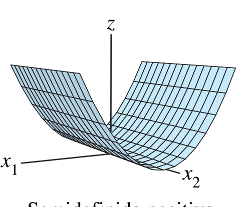
26. Muestre que si una matriz A de $n \times n$ es definida positiva, entonces existe una matriz B definida positiva tal que $A = B^TB$. [*Sugerencia:* Escriba $A = PDP^T$, con $P^T = P^{-1}$. Produzca una matriz diagonal C tal que $D = C^TC$, y sea $B = PCP^T$. Muestre que B funciona.]

27. Sean A y B matrices simétricas de $n \times n$ cuyos valores propios sean todos positivos. Muestre que todos los valores propios de $A + B$ son positivos. [*Sugerencia:* Considere formas cuadráticas.]

28. Sea A una matriz simétrica invertible de $n \times n$. Muestre que si la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ es definida positiva, también lo es la forma cuadrática $\mathbf{x}^TA^{-1}\mathbf{x}$. [*Sugerencia:* Considere los valores propios.]

SG Dominio de la diagonalización y las formas cuadráticas 7 a 8 (Mastering: Diagonalization and Quadratic Forms 7-8)

SOLUCIÓN AL PROBLEMA DE PRÁCTICA



Semidefinida positiva

Realice un cambio ortogonal de variable $\mathbf{x} = P\mathbf{y}$, y escriba

$$\mathbf{x}^TA\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

como en (4). Si un valor propio —por ejemplo, λ_i— fuera negativo, entonces $\mathbf{x}^TA\mathbf{x}$ sería negativa para el $\mathbf{x}$ correspondiente a $\mathbf{y} = \mathbf{e}_i$ (la columna i-ésima de I_n). Así que todos los valores propios de una forma cuadrática semidefinida positiva deben ser no negativos. De manera recíproca, si los valores propios son no negativos, la ampliación anterior muestra que $\mathbf{x}^TA\mathbf{x}$ debe ser semidefinida positiva.

7.3 OPTIMIZACIÓN RESTRINGIDA

A menudo los ingenieros, economistas, científicos y matemáticos necesitan encontrar el valor máximo o mínimo de una forma cuadrática $Q(\mathbf{x})$ para $\mathbf{x}$ en algún conjunto específico. De manera típica, el problema puede plantearse en una forma tal que $\mathbf{x}$ varíe

sobre el conjunto de vectores unitarios. Como se verá más adelante, este *problema de optimización restringida* tiene una solución interesante y elegante. El ejemplo 6 que se presenta enseguida y el análisis de la sección 7.5 ilustran cómo surgen tales problemas en la práctica.

El requisito de que un vector **x** en $\mathbb{R}^n$ sea un vector unitario puede plantearse de varias maneras equivalentes:

$$\|\mathbf{x}\| = 1, \qquad \|\mathbf{x}\|^2 = 1, \qquad \mathbf{x}^T\mathbf{x} = 1$$

y

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 1 \tag{1}$$

Se usará $\mathbf{x}^T\mathbf{x} = 1$, pero la versión ampliada (1) es la que se emplea comúnmente en las aplicaciones.

Cuando una forma cuadrática Q no tiene términos de producto cruzado, es fácil encontrar los valores máximo y mínimo de $Q(\mathbf{x})$ para $\mathbf{x}^T\mathbf{x} = 1$.

EJEMPLO 1 Encuentre los valores máximo y mínimo de $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ sujetos a la restricción de que $\mathbf{x}^T\mathbf{x} = 1$.

Solución Como x_2^2 y x_3^2 son no negativos, observe que

$$4x_2^2 \le 9x_2^2 \qquad \text{y} \qquad 3x_3^2 \le 9x_3^2$$

y, por lo tanto,

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\le 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

siempre que $x_1^2 + x_2^2 + x_3^2 = 1$. Entonces el valor máximo de $Q(\mathbf{x})$ no puede ser mayor que 9 cuando **x** es un vector unitario. Más aún, $Q(\mathbf{x}) = 9$ cuando $\mathbf{x} = (1, 0, 0)$. Entonces 9 es el valor máximo de $Q(\mathbf{x})$ para $\mathbf{x}^T\mathbf{x} = 1$.

Para encontrar el valor mínimo de $Q(\mathbf{x})$, observe que

$$9x_1^2 \ge 3x_1^2, \qquad 4x_2^2 \ge 3x_2^2$$

y, por lo tanto,

$$Q(\mathbf{x}) \ge 3x_1^2 + 3x_2^2 + 3x_3^2 = 3(x_1^2 + x_2^2 + x_3^2) = 3$$

siempre que $x_1^2 + x_2^2 + x_3^2 = 1$. Asimismo, $Q(\mathbf{x}) = 3$ cuando $x_1 = 0$, $x_2 = 0$, y $x_3 = 1$. De manera que 3 es el valor mínimo de $Q(\mathbf{x})$ cuando $\mathbf{x}^T\mathbf{x} = 1$. ■

En el ejemplo 1, es fácil advertir que la matriz de la forma cuadrática Q tiene valores propios 9, 4 y 3, y que los valores propios mayor y menor son iguales, respectivamente, al máximo y el mínimo (restringidos) de $Q(\mathbf{x})$. Lo mismo es válido para cualquier forma cuadrática, como se verá más adelante.

EJEMPLO 2 Sea $A = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$, y para **x** en $\mathbb{R}^2$ sea $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$. La figura 1 muestra la gráfica de Q. En la figura 2 se presenta solamente la porción de la gráfica situada den-

tro de un cilindro; la intersección del cilindro con la superficie es el conjunto de puntos (x_1, x_2, z) tales que $z = Q(x_1, x_2)$ y $x_1^2 + x_2^2 = 1$. Las "alturas" de estos puntos son los valores restringidos de $Q(\mathbf{x})$. Geométricamente, el problema de optimización consiste en localizar los puntos más alto y más bajo de la curva de intersección.

Los dos puntos más altos de la curva están 7 unidades por encima del plano x_1x_2, y ocurren donde $x_1 = 0$ y $x_2 = \pm 1$. Estos puntos corresponden al valor propio 7 de A y a los vectores propios $\mathbf{x} = (0, 1)$ y $-\mathbf{x} = (0, -1)$. De manera similar, los dos puntos más bajos de la curva están 3 unidades por encima del plano x_1x_2, y corresponden al valor propio 3 y a los vectores propios $(1, 0)$ y $(-1, 0)$.

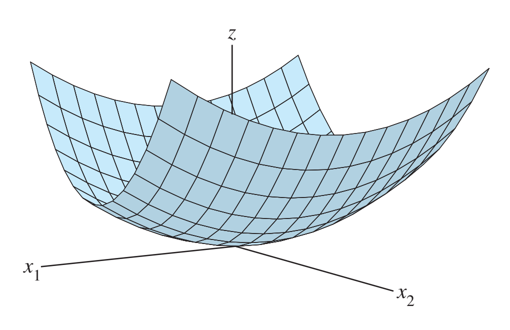


FIGURA 1 $z = 3x_1^2 + 7x_2^2$.

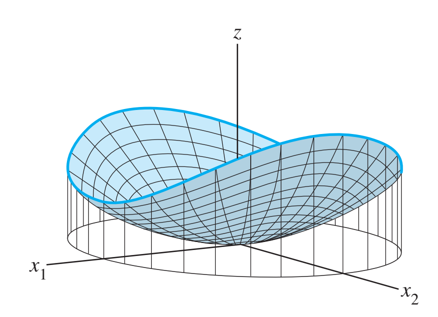


FIGURA 2 La intersección de $z = 3x_1^2 + 7x_2^2$ y el cilindro $x_1^2 + x_2^2 = 1$.

Todo punto sobre la curva de intersección de la figura 2 tiene una coordenada z entre 3 y 7, y para cualquier número t entre 3 y 7 hay un vector unitario $\mathbf{x}$ tal que $Q(\mathbf{x}) = t$. En otras palabras, el conjunto de todos los valores posibles de $\mathbf{x}^TA\mathbf{x}$, para $\|\mathbf{x}\| = 1$, es el intervalo cerrado $3 \leq t \leq 7$.

Puede mostrarse que para cualquier matriz simétrica A, el conjunto de todos los valores posibles de $\mathbf{x}^TA\mathbf{x}$, para $\|\mathbf{x}\| = 1$, es un intervalo cerrado sobre el eje de los reales. (Vea el ejercicio 13.) Denote los extremos izquierdo y derecho de este intervalo mediante m y M, respectivamente. Esto es, sean

$$m = \text{mín}\,\{\mathbf{x}^TA\mathbf{x} : \|\mathbf{x}\| = 1\}, \quad M = \text{máx}\,\{\mathbf{x}^TA\mathbf{x} : \|\mathbf{x}\| = 1\} \tag{2}$$

El ejercicio 12 pide probar que si λ es un valor propio de A, entonces $m \leq \lambda \leq M$. El teorema siguiente postula que m y M son, en sí mismos, valores propios de A, como en el ejemplo 2.[1]

TEOREMA 6

Sea A una matriz simétrica, y defínanse m y M como en (2). Entonces M es el valor propio λ_1 más grande de A, y m es el valor propio más pequeño de A. El valor de $\mathbf{x}^TA\mathbf{x}$ es M cuando $\mathbf{x}$ es un vector propio unitario $\mathbf{u}_1$ correspondiente a M. El valor de $\mathbf{x}^TA\mathbf{x}$ es m cuando $\mathbf{x}$ es un vector propio unitario correspondiente a m.

[1]El uso de los términos *mínimo* y *máximo* en (2), así como de *menor* y *mayor* en el teorema, se refiere al orden natural de los números reales, no a magnitudes.

DEMOSTRACIÓN Diagonalice ortogonalmente a A como PDP^{-1}. Se sabe que

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} \quad \text{cuando } \mathbf{x} = P\mathbf{y} \tag{3}$$

También,

$$\|\mathbf{x}\| = \|P\mathbf{y}\| = \|\mathbf{y}\| \quad \text{para todo } \mathbf{y}$$

porque $P^TP = I$ y $\|P\mathbf{y}\|^2 = (P\mathbf{y})^T(P\mathbf{y}) = \mathbf{y}^TP^TP\mathbf{y} = \mathbf{y}^T\mathbf{y} = \|\mathbf{y}\|^2$. En particular, $\|\mathbf{y}\| = 1$ si, y sólo si, $\|\mathbf{x}\| = 1$. Entonces $\mathbf{x}^TA\mathbf{x}$ y $\mathbf{y}^TD\mathbf{y}$ asumen el mismo conjunto de valores cuando $\mathbf{x}$ e $\mathbf{y}$ recorren el conjunto de todos los vectores unitarios.

Para simplificar la notación, se supondrá que A es una matriz de 3×3 con valores propios $a \geq b \geq c$. Acomode las columnas (vectores propios) de P en forma tal que $P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3]$ y

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

Dado cualquier vector unitario $\mathbf{y}$ en $\mathbb{R}^3$ con coordenadas y_1, y_2, y_3, observe que

$$\begin{aligned} ay_1^2 &= ay_1^2 \\ by_2^2 &\leq ay_2^2 \\ cy_3^2 &\leq ay_3^2 \end{aligned}$$

Al sumar estas desigualdades, se obtiene

$$\begin{aligned} \mathbf{y}^TD\mathbf{y} &= ay_1^2 + by_2^2 + cy_3^2 \\ &\leq ay_1^2 + ay_2^2 + ay_3^2 \\ &= a(y_1^2 + y_2^2 + y_3^2) \\ &= a\|\mathbf{y}\|^2 = a \end{aligned}$$

Entonces $M \leq a$, de acuerdo con la definición de M. Sin embargo, $\mathbf{y}^TD\mathbf{y} = a$ cuando $\mathbf{y} = \mathbf{e}_1 = (1, 0, 0)$, así que, de hecho, $M = a$. Según (3), el $\mathbf{x}$ que corresponde a $\mathbf{y} = \mathbf{e}_1$ es el vector propio $\mathbf{u}_1$ de A, porque

$$\mathbf{x} = P\mathbf{e}_1 = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,]\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{u}_1$$

Entonces $M = a = \mathbf{e}_1^T D\mathbf{e}_1 = \mathbf{u}_1^T A\mathbf{u}_1$, lo cual demuestra el enunciado acerca de M. Un argumento similar muestra que m es el valor propio menor, c, y este valor de $\mathbf{x}^TA\mathbf{x}$ se alcanza cuando $\mathbf{x} = P\mathbf{e}_3 = \mathbf{u}_3$. ■

EJEMPLO 3 Sea $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$. Determine el valor máximo de la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ sujeto a la restricción de que $\mathbf{x}^T\mathbf{x} = 1$, y encuentre un vector unitario en el cual se alcance este valor máximo.

Solución De acuerdo con el teorema 6, buscamos el valor propio mayor de A. La ecuación característica resulta ser

$$0 = -\lambda^3 + 10\lambda^2 - 27\lambda + 18 = -(\lambda - 6)(\lambda - 3)(\lambda - 1)$$

El mayor valor propio es 6.

El máximo restringido de $\mathbf{x}^TA\mathbf{x}$ se alcanza cuando $\mathbf{x}$ es un vector propio para $\lambda = 6$. Al resolver $(A - 6I)\mathbf{x} = 0$, se encuentra un vector propio $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$.

En aplicaciones posteriores será necesario considerar los valores de $\mathbf{x}^TA\mathbf{x}$ cuando $\mathbf{x}$ no sólo es un vector unitario, sino que también es ortogonal al vector propio $\mathbf{u}_1$ mencionado en el teorema 6. Este caso se trata en el teorema siguiente.

TEOREMA 7

Sean A, λ_1 y $\mathbf{u}_1$ como en el teorema 6. Entonces el valor máximo de $\mathbf{x}^TA\mathbf{x}$ sujeto a las restricciones

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0$$

es el segundo valor propio más grande, λ_2, y se alcanza este máximo cuando $\mathbf{x}$ es un vector propio $\mathbf{u}_2$ correspondiente a λ_2.

El teorema 7 puede demostrarse mediante un argumento semejante al anterior en el cual el teorema se redujo al caso en el que la matriz de la forma cuadrática es diagonal. El siguiente ejemplo proporciona una idea de la demostración en el caso de una matriz diagonal.

EJEMPLO 4 Encuentre el valor máximo de $9x_1^2 + 4x_2^2 + 3x_3^2$ sujeto a las restricciones $\mathbf{x}^T\mathbf{x} = 1$ y $\mathbf{x}^T\mathbf{u}_1 = \mathbf{0}$, donde $\mathbf{u}_1 = (1, 0, 0)$. Observe que $\mathbf{u}_1$ es un vector propio unitario correspondiente al valor propio mayor $\lambda = 9$ de la matriz de la forma cuadrática.

Solución Si las coordenadas de $\mathbf{x}$ son x_1, x_2, x_3, entonces la restricción $\mathbf{x}^T\mathbf{u}_1 = 0$ significa simplemente que $x_1 = 0$. Para un vector unitario, $x_2^2 + x_3^2 = 1$, y

$$\begin{aligned} 9x_1^2 + 4x_2^2 + 3x_3^2 &= 4x_2^2 + 3x_3^2 \\ &\leq 4x_2^2 + 4x_3^2 \\ &= 4(x_2^2 + x_3^2) \\ &= 4 \end{aligned}$$

Entonces el máximo restringido de la forma cuadrática no excede a 4. Y este valor se alcanza para $\mathbf{x} = (0, 1, 0)$, el cual es un vector propio para el segundo valor propio mayor de la matriz de la forma cuadrática.

EJEMPLO 5 Sea A la matriz del ejemplo 3 y sea $\mathbf{u}_1$ un vector propio unitario correspondiente al valor propio mayor de A. Encuentre el valor máximo de $\mathbf{x}^TA\mathbf{x}$ sujeto a las condiciones

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0 \tag{4}$$

Solución De acuerdo con el ejemplo 3, el segundo mayor valor propio de A es $\lambda = 3$. Resuelva $(A - 3I)\mathbf{x} = \mathbf{0}$ para encontrar un vector propio y normalícelo para obtener

$$\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

El vector $\mathbf{u}_2$ es, automáticamente, ortogonal a $\mathbf{u}_1$ porque los vectores corresponden a valores propios diferentes. Entonces el máximo de $\mathbf{x}^T A\mathbf{x}$ sujeto a las restricciones de (4) es 3, el cual se alcanza cuando $\mathbf{x} = \mathbf{u}_2$.

El teorema siguiente generaliza el teorema 7 y, junto con el teorema 6, proporciona una caracterización útil de *todos* los valores propios de A. Se omite la demostración.

TEOREMA 8

Sea A una matriz simétrica de $n \times n$ con una diagonalización ortogonal $A = PDP^{-1}$, donde las entradas sobre la diagonal de D están acomodadas para que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ y donde las columnas de P son vectores propios unitarios correspondientes $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Entonces para $k = 2, \ldots, n$, el valor máximo de $\mathbf{x}^T A\mathbf{x}$ sujeto a las restricciones

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0, \quad \ldots, \quad \mathbf{x}^T\mathbf{u}_{k-1} = 0$$

es el valor propio λ_k, y alcanza su máximo en $\mathbf{x} = \mathbf{u}_k$.

El teorema 8 será útil en las secciones 7.4 y 7.5. La siguiente aplicación sólo requiere del teorema 6.

EJEMPLO 6 Durante el próximo año, el gobierno de cierto condado planea reparar x cientos de millas de caminos públicos y puentes, y mejorar y cientos de acres de parques y áreas recreativas. El condado tiene que decidir cómo asignar sus recursos (fondos, equipo, mano de obra, etc.) entre estos dos proyectos. Si resulta más eficiente, por costos, trabajar de manera simultánea en ambos proyectos en lugar de atender solamente uno, entonces x y y podrían satisfacer una *restricción* tal como

$$4x^2 + 9y^2 \leq 36$$

Vea la figura 3. Cada punto (x, y) localizado en el *conjunto factible* sombreado representa una obra pública posible programada para el año. Los puntos ubicados sobre la curva de restricción, $4x^2 + 9y^2 = 36$, utilizan las cantidades máximas de recursos disponibles.

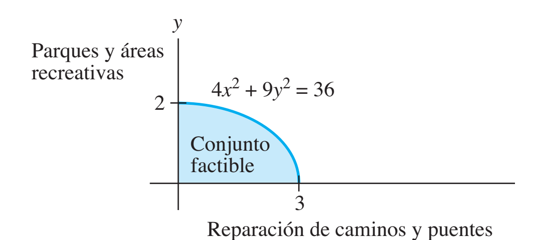


FIGURA 3 Programa de obras públicas.

Al seleccionar su programa de obras públicas, el condado quiere tomar en cuenta la opinión de los residentes. Para medir el valor o *utilidad* que los residentes asignan a los diversos programas de trabajo (x, y), los economistas utilizan a veces una función tal como

$$q(x, y) = xy$$

El conjunto de puntos (x, y) en el cual $q(x, y)$ es una constante se llama *curva de indiferencia.* En la figura 4 se muestran tres curvas de este tipo. Los puntos ubicados a lo largo de la curva de indiferencia corresponden a las alternativas que los residentes del condado, como grupo, encontrarían igualmente valiosas.[2] Encuentre el programa de obras públicas que maximice la función de utilidad q.

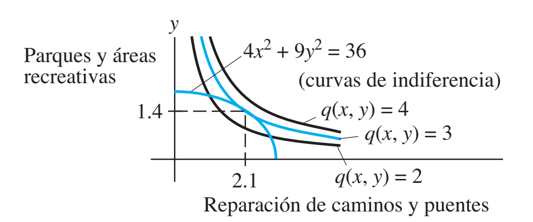


FIGURA 4 El programa de obras públicas óptimo es (2.1, 1.4).

Solución La ecuación de restricción $4x^2 + 9y^2 = 36$ no describe un conjunto de vectores unitarios, pero un cambio de variable puede solucionar ese problema. Reescriba la restricción en forma de

$$\left(\frac{x}{3}\right)^2 + \left(\frac{y}{2}\right)^2 = 1$$

y defina

$$x_1 = \frac{x}{3}, \quad x_2 = \frac{y}{2}, \qquad \text{esto es,} \quad x = 3x_1 \quad \text{y} \quad y = 2x_2$$

Entonces la ecuación de restricción se convierte en

$$x_1^2 + x_2^2 = 1$$

y la función de utilidad se convierte en $q(3x_1, 2x_2) = (3x_1)(2x_2) = 6x_1x_2$. Sea $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.
Entonces el problema consiste en maximizar $Q(\mathbf{x}) = 6x_1x_2$ sujeto a $\mathbf{x}^T\mathbf{x} = 1$. Observe que $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, donde

$$A = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}$$

Los valores propios de A son ± 3, con vectores propios $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ para $\lambda = 3$ y $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ para $\lambda = -3$. Entonces el valor máximo de $Q(\mathbf{x}) = q(x_1, x_2)$ es 3, el cual se alcanza cuando $x_2 = 1/\sqrt{2}$. y $x_1 = 1/\sqrt{2}$.

[2]Las curvas de indiferencia se analizan en Michael D. Intriligator, Ronald G. Bodkin, y Cheng Hsiao, *Econometric Models, Techniques, and Applications* (Upper Saddle River, NJ: Prentice-Hall, 1996).

En términos de las variables originales, el programa de obras públicas óptimo es $x = 3x_1 = 3/\sqrt{2} \approx 2.1$ cientos de millas de caminos y $y = 2x_2 = \sqrt{2} \approx 1.4$ cientos de acres de parques y áreas recreativas. El programa de obras públicas óptimo es el punto donde se encuentran la curva de restricción y la curva de indiferencia $q(x, y) = 3$. Los puntos (x, y) con una utilidad mayor están sobre curvas de indiferencia que no tocan la curva restringida. Vea la figura 4.

PROBLEMAS DE PRÁCTICA

1. Sea $Q(\mathbf{x}) = 3x_1^2 + 3x_2^2 + 2x_1x_2$. Encuentre un cambio de variable que transforma a Q en una forma cuadrática sin términos de producto cruzado, y proporcione la nueva forma cuadrática.
2. Con Q igual que en el problema 1, encuentre el valor máximo de $Q(\mathbf{x})$ sujeto a la restricción de que $\mathbf{x}^T\mathbf{x} = 1$, y encuentre un vector unitario en el que se alcance el máximo.

7.3 EJERCICIOS

En los ejercicios 1 y 2, encuentre el cambio de variable $\mathbf{x} = P\mathbf{y}$ que transforma la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ en $\mathbf{y}^TD\mathbf{y}$ como se muestra.

1. $5x_1^2 + 6x_2^2 + 7x_3^2 + 4x_1x_2 - 4x_2x_3 = 9y_1^2 + 6y_2^2 + 3y_3^2$
2. $3x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3 = 5y_1^2 + 2y_2^2$
[*Sugerencia:* **x** y **y** deben tener el mismo número de coordenadas, así que la forma cuadrática aquí mostrada debe tener un coeficiente de cero para y_2^3.]

En los ejercicios 3 a 6, encuentre (a) el valor máximo de $Q(\mathbf{x})$ sujeto a la restricción $\mathbf{x}^T\mathbf{x} = 1$, (b) un vector unitario **u** donde se alcance este máximo, y (c) el máximo de $Q(\mathbf{x})$ sujeto a las restricciones $\mathbf{x}^T\mathbf{x} = 1$ y $\mathbf{x}^T\mathbf{u} = 0$.

3. $Q(\mathbf{x}) = 5x_1^2 + 6x_2^2 + 7x_3^2 + 4x_1x_2 - 4x_2x_3$ (Vea el ejercicio 1.)
4. $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$ (Vea el ejercicio 2.)
5. $Q(\mathbf{x}) = 5x_1^2 + 5x_2^2 - 4x_1x_2$
6. $Q(\mathbf{x}) = 7x_1^2 + 3x_2^2 + 3x_1x_2$
7. Sea $Q(\mathbf{x}) = -2x_1{}^2 - x_2{}^2 + 4x_1x_2 + 4x_2x_3$. Encuentre un vector unitario **x** en $\mathbb{R}^3$ en el cual se maximice $Q(\mathbf{x})$, sujeto a $\mathbf{x}^T\mathbf{x} = 1$. [*Sugerencia:* Los valores propios de la matriz de la forma cuadrática Q son 2, -1 y -4.]
8. Sea $Q(\mathbf{x}) = 7x_1^2 + x_2^2 + 7x_3^2 - 8x_1x_2 - 4x_1x_3 - 8x_2x_3$. Encuentre un vector unitario **x** de $\mathbb{R}^3$ en el cual $Q(\mathbf{x})$ se maximice, sujeto a $\mathbf{x}^T\mathbf{x} = 1$. [*Sugerencia:* Los valores propios de la matriz de la forma cuadrática Q son 9 y -3.]
9. Encuentre el valor máximo de $Q(\mathbf{x}) = 7x_1^2 + 3x_2^2 - 2x_1x_2$, sujeto a la restricción $x_1^2 + x_2^2 = 1$. (No continúe sino hasta encontrar un vector en el que se alcance el máximo.)
10. Encuentre el valor máximo de $Q(\mathbf{x}) = -3x_1^2 + 5x_2^2 - 2x_1x_2$, sujeto a la restricción $x_1^2 + x_2^2 = 1$. (No continúe sino hasta encontrar un vector en el que se alcance el máximo.)
11. Suponga que **x** es un vector propio unitario de una matriz A correspondiente a un valor propio 3. ¿Cuál es el valor de $\mathbf{x}^TA\mathbf{x}$?
12. Sea λ cualquier valor propio de una matriz simétrica A. Justifique el enunciado emitido en esta sección acerca de que $m \leq \lambda \leq M$, donde m y M están definidas como en (2). [*Sugerencia:* Encuentre un **x** tal que $\lambda = \mathbf{x}^TA\mathbf{x}$.]
13. Sea A una matriz simétrica de $n \times n$, denote con M y m los valores máximo y mínimo de la forma cuadrática $\mathbf{x}^TA\mathbf{x}$, y denote los vectores propios unitarios correspondientes por medio de $\mathbf{u}_1$ y $\mathbf{u}_n$. Los cálculos siguientes muestran que, dado cualquier número t entre M y m, existe un vector unitario **x** tal que $t = \mathbf{x}^TA\mathbf{x}$. Verifique si $t = (1 - \alpha)m + \alpha M$ para algún número α entre 0 y 1. Luego haga $\mathbf{x} = \sqrt{1-\alpha}\mathbf{u}_n + \sqrt{\alpha}\mathbf{u}_1$, y muestre que $\mathbf{x}^T\mathbf{x} = 1$ y $\mathbf{x}^TA\mathbf{x} = t$.

[M] En los ejercicios 14 a 17, siga las instrucciones dadas para los ejercicios 3 a 6.

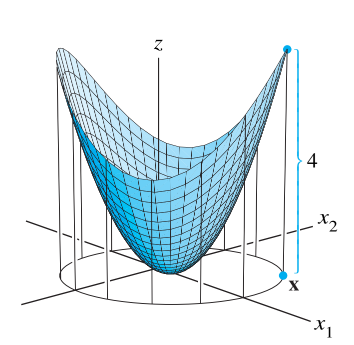
14. $x_1x_2 + 3x_1x_3 + 30x_1x_4 + 30x_2x_3 + 3x_2x_4 + x_3x_4$

15. $3x_1x_2 + 5x_1x_3 + 7x_1x_4 + 7x_2x_3 + 5x_2x_4 + 3x_3x_4$

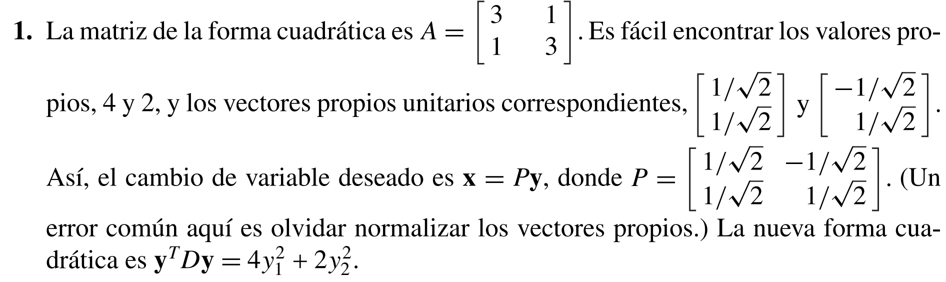
16. $4x_1^2 - 6x_1x_2 - 10x_1x_3 - 10x_1x_4 - 6x_2x_3 - 6x_2x_4 - 2x_3x_4$

17. $-6x_1^2 - 10x_2^2 - 13x_3^2 - 13x_4^2 - 4x_1x_2 - 4x_1x_3 - 4x_1x_4 + 6x_3x_4$

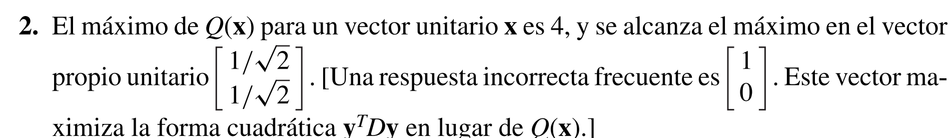
SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA



El máximo valor de $Q(\mathbf{x})$ sujeto a $\mathbf{x}^T\mathbf{x} = 1$ es 4.



1. La matriz de la forma cuadrática es $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Es fácil encontrar los valores propios, 4 y 2, y los vectores propios unitarios correspondientes, $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ y $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. Así, el cambio de variable deseado es $\mathbf{x} = P\mathbf{y}$, donde $P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. (Un error común aquí es olvidar normalizar los vectores propios.) La nueva forma cuadrática es $\mathbf{y}^T D\mathbf{y} = 4y_1^2 + 2y_2^2$.



2. El máximo de $Q(\mathbf{x})$ para un vector unitario $\mathbf{x}$ es 4, y se alcanza el máximo en el vector propio unitario $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. [Una respuesta incorrecta frecuente es $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Este vector maximiza la forma cuadrática $\mathbf{y}^T D\mathbf{y}$ en lugar de $Q(\mathbf{x})$.]

7.4 LA DESCOMPOSICIÓN EN VALORES SINGULARES

Los teoremas de diagonalización presentados en las secciones 5.3 y 7.1 forman parte de muchas aplicaciones interesantes. Por desgracia, como se sabe, no todas las matrices pueden factorizarse como $A = PDP^{-1}$ con diagonal D. Sin embargo, ¡es posible efectuar una factorización $A = QDP^{-1}$ para cualquier matriz A de $m \times n$! Una factorización especial de este tipo, llamada *descomposición en valores singulares*, es una de las factorizaciones de matrices más útiles que existen en el álgebra lineal aplicada.

La descomposición en valores singulares se basa en la siguiente propiedad de la diagonalización ordinaria que se puede imitar para aplicarla en matrices rectangulares: Los valores absolutos de los valores propios de una matriz A simétrica miden las cantidades en que A estira o reduce ciertos vectores (los vectores propios). Si $A\mathbf{x} = \lambda\mathbf{x}$ y $\|\mathbf{x}\| = 1$, entonces

$$\|A\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\| = |\lambda| \tag{1}$$

Si λ_1 es el valor propio con la mayor magnitud, entonces un vector propio unitario correspondiente $\mathbf{v}_1$ identifica una dirección en la cual el efecto de estiramiento de A es el mayor. Esto es, la longitud de $A\mathbf{x}$ se maximiza cuando $\mathbf{x} = \mathbf{v}_1$, y $\|A\mathbf{v}_1\| = |\lambda_1|$, de acuerdo con (1). Esta descripción de $\mathbf{v}_1$ y $|\lambda_1|$ tiene un análogo para matrices rectangulares que conducirá a la descomposición en valores singulares.

EJEMPLO 1 Si $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea la esfera unitaria $\{\mathbf{x} : \|\mathbf{x}\| = 1\}$ en $\mathbb{R}^3$ sobre una elipse en $\mathbb{R}^2$, lo cual se muestra en la figura 1. Encuentre un vector unitario $\mathbf{x}$ donde se maximice la longitud de $\|A\mathbf{x}\|$, y calcule esta longitud máxima.

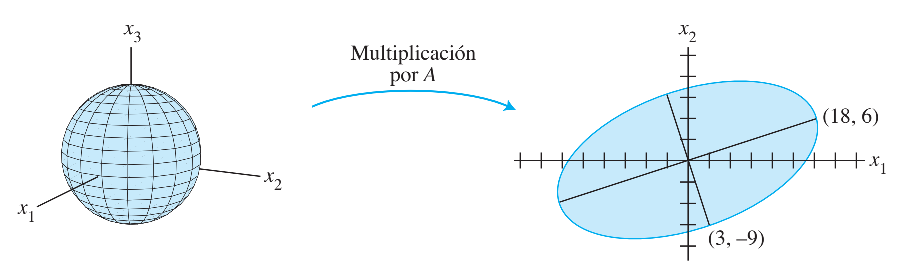


FIGURA 1 Una transformación de $\mathbb{R}^3$ a $\mathbb{R}^2$.

Solución La cantidad $\|A\mathbf{x}\|^2$ se maximiza en el mismo $\mathbf{x}$ que maximiza $\|A\mathbf{x}\|$, y $\|A\mathbf{x}\|^2$ es más fácil de estudiar. Observe que

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^T A^T A\mathbf{x} = \mathbf{x}^T(A^T A)\mathbf{x}$$

También, A^TA es una matriz simétrica, puesto que $(A^TA)^T = A^TA^{TT} = A^TA$. Así que el problema ahora es maximizar la forma cuadrática $\mathbf{x}^T(A^TA)\mathbf{x}$ sujeta a la restricción $\|\mathbf{x}\| = 1$. Éste es un problema de la sección 7.3 y se conoce la solución. De acuerdo con el teorema 6, el valor máximo es el valor propio λ_1 más grande de A^TA. Asimismo, se alcanza el valor máximo en un vector propio unitario de A^TA correspondiente a λ_1.

Para la matriz A dada en este ejemplo,

$$A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$$

Los valores propios de A^TA son $\lambda_1 = 360$, $\lambda_2 = 90$, y $\lambda_3 = 0$. Los vectores propios unitarios correspondientes son, respectivamente.

$$\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

El valor máximo de $\|A\mathbf{x}\|^2$ es 360, que se alcanza cuando $\mathbf{x}$ es el vector unitario $\mathbf{v}_1$. El vector $A\mathbf{v}_1$ es el punto ubicado sobre la elipse de la figura 1 que está más alejado del origen, a saber,

$$A\mathbf{v}_1 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 18 \\ 6 \end{bmatrix}$$

Para $\|\mathbf{x}\| = 1$, el valor máximo de $\|A\mathbf{x}\|$ es $\|A\mathbf{v}_1\| = \sqrt{360} = 6\sqrt{10}$.

El ejemplo 1 sugiere que el efecto de A sobre la esfera unitaria en $\mathbb{R}^3$ está relacionado con la forma cuadrática $\mathbf{x}^T(A^TA)\mathbf{x}$. De hecho, todo el comportamiento geométrico de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ está ligado con esta forma cuadrática, como se verá más adelante.

Los valores singulares de una matriz de $m \times n$

Sea A una matriz de $m \times n$. Entonces A^TA es simétrica y puede diagonalizarse ortogonalmente. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base ortonormal para $\mathbb{R}^n$ que consiste en los vectores propios de A^TA, y sean $\lambda_1, \ldots, \lambda_n$ los valores propios de A^TA asociados. Entonces, para $1 \leq i \leq n$,

$$\begin{aligned} \|A\mathbf{v}_i\|^2 &= (A\mathbf{v}_i)^T A\mathbf{v}_i = \mathbf{v}_i^T A^T A\mathbf{v}_i \\ &= \mathbf{v}_i^T(\lambda_i \mathbf{v}_i) \qquad \text{Puesto que } \mathbf{v}_i \text{ es un vector propio de } A^TA \\ &= \lambda_i \qquad \text{Puesto que } \mathbf{v}_i \text{ es un vector unitario} \end{aligned} \tag{2}$$

Así que todos los valores propios de A^TA son no negativos. Al reenumerar, si es necesario, puede suponerse que los valores propios están acomodados de manera que

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

Los **valores singulares** de A son las raíces cuadradas de los valores propios de A^TA, denotados mediante $\sigma_1, \ldots, \sigma_n$, y están acomodados en orden descendente. Esto es, $\sigma_i = \sqrt{\lambda_i}$ para $1 \leq i \leq n$. De acuerdo con (2), *los valores singulares de A son las longitudes de los vectores* $A\mathbf{v}_1, \ldots, A\mathbf{v}_n$.

EJEMPLO 2 Sea A la matriz del ejemplo 1. Como los valores propios de A^TA son 360, 90 y 0, los valores singulares de A son

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \quad \sigma_2 = \sqrt{90} = 3\sqrt{10}, \quad \sigma_3 = 0$$

A partir del ejemplo 1, el primer valor singular de A es el máximo de $\|A\mathbf{x}\|$ sobre todos los vectores unitarios, y se alcanza el máximo en el vector propio unitario $\mathbf{v}_1$. El teorema 7 de la sección 7.3 muestra que el segundo valor singular de A es el máximo de $\|A\mathbf{x}\|$ sobre todos los vectores unitarios que sean *ortogonales a* $\mathbf{v}_1$, y se alcanza este máximo en el segundo vector propio unitario, $\mathbf{v}_2$ (ejercicio 22). Para el $\mathbf{v}_2$ del ejemplo 1,

$$A\mathbf{v}_2 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \end{bmatrix}$$

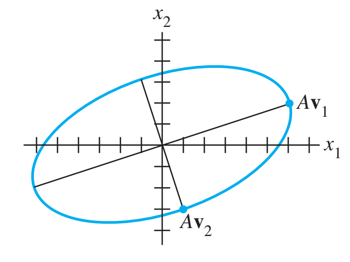


FIGURA 2

Este punto está sobre el eje menor de la elipse de la figura 1, del mismo modo que $A\mathbf{v}_1$ está sobre el eje mayor. (Vea la figura 2.) Los primeros dos valores singulares de A son las longitudes de los semiejes mayor y menor de la elipse.

El hecho de que $A\mathbf{v}_1$ y $A\mathbf{v}_2$ sean ortogonales en la figura 2 no es accidental, como lo muestra el teorema siguiente.

TEOREMA 9 Suponga que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base ortonormal de $\mathbb{R}^n$ que consiste de vectores propios de A^TA, acomodados de manera que los valores propios correspondientes de A^TA satisfagan $\lambda_1 \geq \cdots \geq \lambda_n$, y suponga que A tiene r valores singulares diferentes de cero. Entonces $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ es una base ortogonal para Col A, y rango $A = r$.

DEMOSTRACIÓN Como $\mathbf{v}_i$ y $\lambda_j\mathbf{v}_j$ son ortogonales para $i \neq j$,

$$(A\mathbf{v}_i)^T(A\mathbf{v}_j) = \mathbf{v}_i^T A^T A\mathbf{v}_j = \mathbf{v}_i^T(\lambda_j\mathbf{v}_j) = 0$$

Entonces $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$ es un conjunto ortogonal. Más aún, como las longitudes de los vectores $A\mathbf{v}_1, \ldots, A\mathbf{v}_n$ son los valores singulares de A, y puesto que existen r valores singulares diferentes de cero, $A\mathbf{v}_i \neq \mathbf{0}$ si, y sólo si, $1 \leq i \leq r$. Así que $A\mathbf{v}_1, \ldots, A\mathbf{v}_r$ son vectores linealmente independientes, y están en Col A. Finalmente, para cualquier $\mathbf{y}$ en Col A —por ejemplo, $\mathbf{y} = A\mathbf{x}$— puede escribirse $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, y

$$\begin{aligned}\mathbf{y} = A\mathbf{x} &= c_1A\mathbf{v}_1 + \cdots + c_rA\mathbf{v}_r + c_{r+1}A\mathbf{v}_{r+1} + \cdots + c_nA\mathbf{v}_n \\ &= c_1A\mathbf{v}_1 + \cdots + c_rA\mathbf{v}_r + 0 + \cdots + 0\end{aligned}$$

Entonces $\mathbf{y}$ está en Gen$\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$, lo cual demuestra que $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ es una base (ortogonal) para Col A. Por lo tanto, rango A = dim Col $A = r$. ■

NOTA NUMÉRICA

En algunos casos, el rango de A puede ser muy sensible a pequeños cambios en las entradas de A. El método obvio de contar el número de columnas pivote de A no funciona bien si A se reduce por filas mediante una computadora. A menudo el error de redondeo crea una forma escalonada con rango pleno.

En la práctica, la manera más confiable de estimar el rango de una matriz A grande es contar el número de valores singulares diferentes de cero. En este caso, los valores singulares diferentes de cero extremadamente pequeños se toman como cero para todo fin práctico, y el *rango efectivo* de la matriz es el número que se obtiene al contar los valores singulares diferentes de cero restantes.[1]

La descomposición en valores singulares

La descomposición de A implica una matriz Σ "diagonal" de $m \times n$ de la forma

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} \\ \leftarrow m - r \text{ filas} \end{matrix} \qquad (3)$$

$\uparrow$ $n - r$ columnas

donde D es una matriz diagonal de $r \times r$ para alguna r que no exceda el valor más pequeño de m y n. (Si r es igual a m o a n, o a ambas, alguna de las matrices cero no aparecen.)

[1]En general, la estimación del rango no es un problema simple. Si desea efectuar un análisis de los problemas que implica, vea Philip E. Gill, Walter Murray, y Margaret H. Wright, *Numerical Linear Algebra and Optimization*, vol. 1 (Redwood City, CA: Addison-Wesley, 1991), Sec. 5.8.

TEOREMA 10 **La descomposición en valores singulares**

Sea A una matriz de $m \times n$ con rango r. Entonces existe una matriz Σ de $m \times n$ como la de (3) para la que las entradas diagonales de D son los r primeros valores singulares de A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, y existen una matriz U ortogonal de $m \times m$ y una matriz V ortogonal de $n \times n$ tales que

$$A = U\Sigma V^T$$

Cualquier factorización $A = U\Sigma V^T$, con U y V ortogonales, Σ como en (3), y entradas diagonales positivas en D, es una **descomposición en valores singulares** (o **DVS**) de A. Las matrices U y V no están determinadas en forma única por A, pero las entradas diagonales de Σ son necesariamente los valores singulares de A. Vea el ejercicio 19. Las columnas de U incluidas en una descomposición de este tipo se llaman **vectores singulares izquierdos** de A, y las columnas de V se denominan **vectores singulares derechos** de A.

DEMOSTRACIÓN Sean λ_i y $\mathbf{v}_i$ como en el teorema 9, de modo que $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ es una base ortogonal para Col A. Normalice cada $A\mathbf{v}_i$ para obtener una base ortonormal $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$, donde

$$\mathbf{u}_i = \frac{1}{\|A\mathbf{v}_i\|}A\mathbf{v}_i = \frac{1}{\sigma_i}A\mathbf{v}_i$$

y

$$A\mathbf{v}_i = \sigma_i\mathbf{u}_i \qquad (1 \leq i \leq r) \tag{4}$$

Ahora extienda $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ hasta una base ortonormal $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ de $\mathbb{R}^m$, y sea

$$U = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m\,] \quad \text{y} \quad V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n\,]$$

Por estructuración, U y V son matrices ortogonales. También, de (4),

$$AV = [\,A\mathbf{v}_1 \quad \cdots \quad A\mathbf{v}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,] = [\,\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,]$$

Sea D la matriz diagonal con entradas diagonales $\sigma_1, \ldots, \sigma_r$, y sea Σ como en la ecuación (3) anterior. Entonces

$$\begin{aligned} U\Sigma &= [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m\,]\left[\begin{array}{cccc|c} \sigma_1 & & & 0 & \\ & \sigma_2 & & & 0 \\ & & \ddots & & \\ 0 & & & \sigma_r & \\ \hline & & 0 & & 0 \end{array}\right] \\ &= [\,\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,] \\ &= AV \end{aligned}$$

Como V es una matriz ortogonal, $U\Sigma V^T = AVV^T = A$. ■

Los siguientes dos ejemplos enfocan su atención en la estructura interna de una descomposición en valores singulares. Un algoritmo eficiente y numéricamente estable para esta descomposición usaría un enfoque diferente. Vea la nota numérica que aparece al final de esta sección.

EJEMPLO 3 Use los resultados de los ejemplos 1 y 2 para estructurar una descomposición en valores singulares de $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$.

Solución Una construcción puede dividirse en tres pasos.

SG Cálculo de una DVS 7 a 11 (Computing an SVD 7-11)

Paso 1. Encontrar una diagonalización de A^TA. Esto es, encontrar los valores propios de A^TA y un conjunto ortonormal correspondiente de vectores propios. Si A tiene sólo dos columnas, los cálculos pueden hacerse a mano. Por lo general, las matrices más grandes requieren un programa de matrices. Sin embargo, para la matriz A que se presenta aquí, los datos propios para A^TA se obtienen del ejemplo 1.

Paso 2. Establecer V y Σ. Acomodar los valores propios de A^TA en orden descendente. En el ejemplo 1, los valores propios ya están enlistados en orden descendente: 360, 90 y 0. Los vectores propios unitarios correspondientes, $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, son los vectores singulares derechos de A. Usando el ejemplo 1, se estructura

$$V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}$$

Las raíces cuadradas de los valores propios son los valores singulares:

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

Los valores singulares distintos de cero son las entradas diagonales de D. La matriz Σ es del mismo tamaño que A, con D en la esquina superior izquierda y ceros en las demás posiciones.

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}, \qquad \Sigma = [\,D \quad 0\,] = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

Paso 3. Estructurar U. Cuando A tiene rango r, las primeras r columnas de U son los vectores normalizados obtenidos de $A\mathbf{v}_1, \ldots, A\mathbf{v}_r$. En este ejemplo, A tiene dos valores singulares diferentes de cero, por lo tanto, rango $A = 2$. De la ecuación (2) y del párrafo anterior al ejemplo 2, recuerde que $\|A\mathbf{v}_1\| = \sigma_1$ y $\|A\mathbf{v}_2\| = \sigma_2$. Entonces

$$\mathbf{u}_1 = \frac{1}{\sigma_1}A\mathbf{v}_1 = \frac{1}{6\sqrt{10}}\begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2}A\mathbf{v}_2 = \frac{1}{3\sqrt{10}}\begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$

Observe que $\{\mathbf{u}_1, \mathbf{u}_2\}$ ya es una base para $\mathbb{R}^2$. Entonces no se necesitan vectores adicionales para U, y $U = [\mathbf{u}_1 \quad \mathbf{u}_2]$. La descomposición en valores singulares de A es

---

Consulte la *Guía de estudio* (*Study guide*) para ver comandos de software y graficación en calculadora. Por ejemplo, MATLAB puede producir tanto valores propios como los vectores propios con un comando `erg`.

$$A = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

$$\uparrow \qquad\qquad \uparrow \qquad\qquad \uparrow$$

$$U \qquad\qquad \Sigma \qquad\qquad V^T$$

EJEMPLO 4 Encuentre una descomposición en valores singulares de $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

Solución Primero, calcule $A^TA = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$. Los valores propios de A^TA son 18 y 0, con vectores propios unitarios correspondientes

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Estos vectores unitarios forman las columnas de V:

$$V = [\,\mathbf{v}_1 \quad \mathbf{v}_2\,] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Los valores singulares son $\sigma_1 = \sqrt{18} = 3\sqrt{2}$ y $\sigma_2 = 0$. Como existe únicamente un valor singular diferente de cero, la "matriz" D puede escribirse como un solo número. Esto es, $D = 3\sqrt{2}$. La matriz Σ es del mismo tamaño que A, con D en la esquina superior izquierda.

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

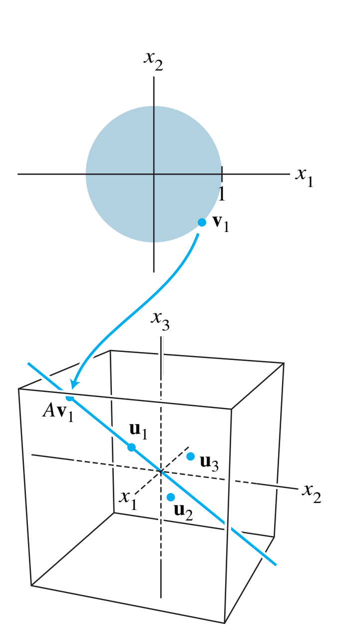


FIGURA 3

Para construir U, primero construya $A\mathbf{v}_1$ y $A\mathbf{v}_2$:

$$A\mathbf{v}_1 = \begin{bmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix}, \quad A\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Como una comprobación de los cálculos, verifique si $\|A\mathbf{v}_1\| = \sigma_1 = 3\sqrt{2}$. Desde luego, $A\mathbf{v}_2 = \mathbf{0}$ porque $\|A\mathbf{v}_2\| = \sigma_2 = 0$. La única columna para U encontrada hasta ahora es

$$\mathbf{u}_1 = \frac{1}{3\sqrt{2}} A\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

Las otras columnas de U se encuentran al ampliar el conjunto $\{\mathbf{u}_1\}$ hasta una base ortonormal para $\mathbb{R}^3$. En este caso, se necesitan dos vectores unitarios ortogonales $\mathbf{u}_2$ y $\mathbf{u}_3$ que sean ortogonales a $\mathbf{u}_1$. (Vea la figura 3.) Cada vector debe satisfacer $\mathbf{u}_1^T\mathbf{x} = 0$, lo cual es equivalente a la ecuación $x_1 - 2x_2 + 2x_3 = 0$. Una base para el conjunto solución de esta ecuación es

$$\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

(Compruebe que $\mathbf{w}_1$ y $\mathbf{w}_2$ son ortogonales a $\mathbf{u}_1$.) Si se aplica el proceso Gram-Schmidt (con normalizaciones) a $\{\mathbf{w}_1, \mathbf{w}_2\}$, se obtiene

$$\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$$

Por último, establezca $U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3]$, tome Σ y V^T de las ecuaciones anteriores, y escriba

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Aplicaciones de la descomposición en valores singulares

La DVS se utiliza a menudo para efectuar una estimación del rango de una matriz, como se observó anteriormente. A continuación se describen de modo breve algunas otras aplicaciones numéricas, y en la sección 7.5 se presenta una aplicación al procesamiento de imágenes.

EJEMPLO 5 (*El número de condición.*) La mayor parte de los cálculos numéricos en los que interviene una ecuación $A\mathbf{x} = \mathbf{b}$ son tan confiables como es posible cuando se usa la DVS de A. Las dos matrices ortogonales U y V no afectan ni la longitud de los vectores ni los ángulos entre éstos (teorema 7 de la sección 6.2). Cualesquiera posibles inestabilidades en los cálculos numéricos se identifican en Σ. Si los valores singulares de A son extremadamente grandes o pequeños, los errores de redondeo son casi inevitables, pero conocer las entradas de Σ y V ayuda a efectuar un análisis del error.

Si A es una matriz invertible de $n \times n$, entonces la razón σ_1/σ_n de los valores singulares mayor y menor proporciona el **número de condición** de A. Los ejercicios 41, 42 y 43 de la sección 2.3 mostraron cómo el número de condición afecta la sensibilidad de una solución de $A\mathbf{x} = \mathbf{b}$ a los cambios (o errores) en las entradas de A. (En realidad, hay varias maneras de calcular un "número de condición" de A, pero la definición que se da aquí se utiliza ampliamente para estudiar $A\mathbf{x} = \mathbf{b}$.)

EJEMPLO 6 (*Bases para los subespacios fundamentales*.) Dada una DVS para una matriz A de $m \times n$, sean $\mathbf{u}_1, \ldots, \mathbf{u}_m$ los vectores singulares izquierdos, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ los vectores singulares derechos, y $\sigma_1, \ldots, \sigma_n$ los valores singulares, y sea r el rango de A. Según el teorema 9,

$$\{\mathbf{u}_1, \ldots, \mathbf{u}_r\} \tag{5}$$

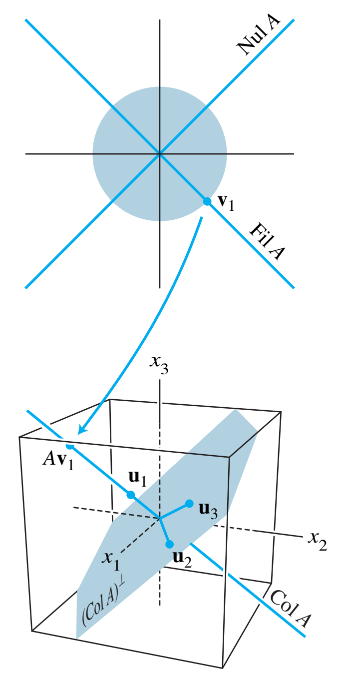
es una base ortonormal para Col A.

Recuerde del teorema 3 de la sección 6.1 que $(\text{Col } A)^\perp = \text{Nul } A^T$. De aquí que,

$$\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\} \tag{6}$$

sea una base ortonormal para Nul A^T.

Como $\|A\mathbf{v}_i\| = \sigma_i$ para $1 \le i \le n$, y σ_i es 0 si, y sólo si, $i > r$, los vectores $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ generan un subespacio de Nul A de dimensión $n - r$. De acuerdo con el teorema del



Los espacios fundamentales en el ejemplo 4.

rango, dim Nul $A = n -$ rango A. Se deduce que

$$\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\} \tag{7}$$

es una base ortonormal para Nul A, según el teorema de la base (de la sección 4.5).

De (5) y (6), el complemento ortogonal de Nul A^T es Col A. Al intercambiar A y A^T, se tiene que $(\text{Nul } A)^{\perp} = \text{Col } A^T = \text{Fil } A$. Por lo tanto, de (7),

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \tag{8}$$

es una base ortonormal para Fil A.

En la figura 4 se resumen las ecuaciones (5) a (8), pero se muestra la base ortogonal $\{\sigma_1\mathbf{u}_1, \ldots, \sigma_r\mathbf{u}_r\}$ para Col A en lugar de la base normalizada, para recordar que $A\mathbf{v}_i = \sigma_i\mathbf{u}_i$ para $1 \leq i \leq r$. Las bases ortonormales explícitas para los cuatro subespacios fundamentales determinados por A son útiles en algunos cálculos, particularmente en problemas de optimización restringida.

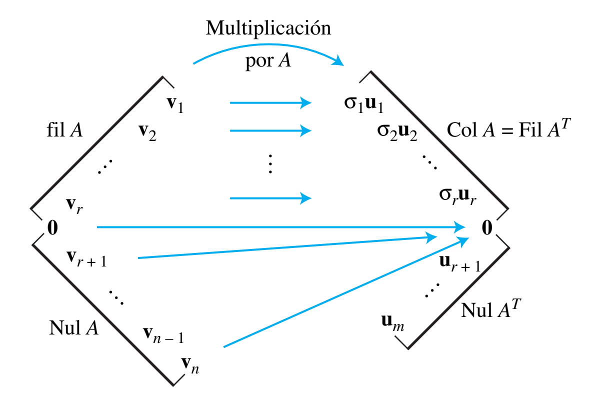


FIGURA 4 Los cuatro subespacios fundamentales y la acción de A.

Los cuatro subespacios fundamentales y el concepto de valores singulares proporcionan los enunciados finales del teorema de la matriz invertible. (Recuerde que los enunciados acerca de A^T se han omitido del teorema, para evitar casi duplicar el número de afirmaciones.) Los otros enunciados se presentaron en las secciones 2.3, 2.9, 3.2, 4.6 y 5.2.

TEOREMA

Teorema de la matriz invertible (terminado)

Sea A una matriz de $n \times n$. Entonces cada uno de los siguientes enunciados es equivalente a la afirmación de que A es una matriz invertible.

u. $(\text{Col } A)^{\perp} = \{\mathbf{0}\}$.

v. $(\text{Nul } A)^{\perp} = \mathbb{R}^n$.

w. Fil $A = \mathbb{R}^n$.

x. A tiene n valores singulares diferentes de cero.

EJEMPLO 7 (*DVS reducida y la seudoinversa de A.*) Cuando Σ contiene filas o columnas de ceros, es posible obtener una descomposición de A más compacta. Usando la notación establecida antes, sea r = rango A, y divida U y V en submatrices cuyos primeros bloques contengan r columnas:

$$\begin{aligned} U &= [\, U_r \quad U_{m-r} \,], \quad \text{donde } U_r = [\, \mathbf{u}_1 \quad \cdots \quad \mathbf{u}_r \,] \\ V &= [\, V_r \quad V_{n-r} \,], \quad \text{donde } V_r = [\, \mathbf{v}_1 \quad \cdots \quad \mathbf{v}_r \,] \end{aligned}$$

Entonces U_r es de $m \times r$ y V_r es de $n \times r$. (Para simplificar la notación, se considera U_{m-r} o V_{n-r} aunque puede ser que alguna de ellas no tenga columnas.) Entonces la multiplicación de matrices partidas muestra que

$$A = [\, U_r \quad U_{m-r} \,] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix} = U_r D V_r^T \tag{9}$$

Esta factorización de A se llama **descomposición en valores singulares reducida** de A. Como las entradas diagonales de D son diferentes de cero, puede formarse la siguiente matriz, llamada **seudoinversa** (también, **inversa Moore-Penrose)** de A:

$$A^+ = V_r D^{-1} U_r^T \tag{10}$$

Los ejercicios suplementarios 12, 13 y 14 incluidos al final del capítulo exploran algunas de las propiedades de la descomposición en valores singulares reducida y de la seudoinversa.

EJEMPLO 8 (*Solución por mínimos cuadrados.*) Dada la ecuación $A\mathbf{x} = \mathbf{b}$, use la seudoinversa de A presentada en (10) para definir

$$\hat{\mathbf{x}} = A^+\mathbf{b} = V_r D^{-1} U_r^T \mathbf{b}$$

Luego, a partir de la DVS de (9),

$$\begin{aligned} A\hat{\mathbf{x}} &= (U_r D V_r^T)(V_r D^{-1} U_r^T \mathbf{b}) \\ &= U_r D D^{-1} U_r^T \mathbf{b} \qquad \text{Porque } V_r^T V_r = I_r \\ &= U_r U_r^T \mathbf{b} \end{aligned}$$

De (5), se deduce que $U_r D_r^T \mathbf{b}$ es la proyección ortogonal $\hat{\mathbf{b}}$ de $\mathbf{b}$ sobre Col A. (Vea el teorema 10 de la sección 6.3.) Entonces $\hat{\mathbf{x}}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. De hecho, esta $\hat{\mathbf{x}}$ tiene la menor longitud de todas las soluciones por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. Vea el ejercicio suplementario 14.

NOTA NUMÉRICA

Los ejemplos 1 a 4 y los ejercicios ilustran el concepto de valores singulares y sugieren cómo realizar cálculos a mano. En la práctica, se debería evitar el cálculo de A^TA, puesto que cualesquiera errores en las entradas de A se elevan al cuadrado en las entradas de A^TA. Existen métodos iterativos rápidos que producen los valores singulares y vectores singulares de A con precisión de hasta muchas posiciones decimales.

Lecturas adicionales

Horn, Roger A. y Charles R. Johnson, *Matrix Analysis*, vol. 1 (Cambridge: Cambridge University Press, 1985), pp. 414-445.

Long, Cliff, "Visualization of Matrix Singular Value Decomposition, *Mathematics Magazine* **56** (1983), pp. 161-167.

Moler, C. B. y D. Morrison. "Singular Value Analysis of Cryptograms." *Amer. Math. Monthly* **90** (1983), pp. 78-87.

Strang, Gilbert, *Linear Algebra and Its Applications*, 3a. ed. (San Diego: Harcourt Brace Jovanovich, 1988), pp. 442-452.

Watkins, David S., *Fundamentals of Matrix Computations* (Nueva York: Wiley, 1991), pp 390-398, 409-421.

PROBLEMA DE PRÁCTICA

Dada una descomposición en valores singulares, $A = U\Sigma V^T$, encuentre una DVS para A^T. ¿Cómo están relacionados los valores singulares de A y de A^T?

7.4 EJERCICIOS

Encuentre los valores singulares de las matrices de los ejercicios 1 a 4.

1. $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$

2. $\begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} \sqrt{6} & 1 \\ 0 & \sqrt{6} \end{bmatrix}$

4. $\begin{bmatrix} \sqrt{3} & 2 \\ 0 & \sqrt{3} \end{bmatrix}$

Encuentre una DVS de cada matriz de los ejercicios 5 a 12. [*Sugerencia:* En el ejercicio 11, una opción para U es $\begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$. En el ejercicio 12, una columna de U puede ser $\begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$.

5. $\begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}$

6. $\begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$

7. $\begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$

8. $\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$

9. $\begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$

10. $\begin{bmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{bmatrix}$

11. $\begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$

13. Encuentre la DVS de $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$. [*Sugerencia:* Trabaje con A^T.]

14. En el ejercicio 7, encuentre un vector unitario $\mathbf{x}$ en el cual $A\mathbf{x}$ tenga su longitud máxima.

15. Suponga que la siguiente factorización es una DVS para una matriz A, con las entradas de U y V redondeadas a dos posiciones decimales.

$$A = \begin{bmatrix} .40 & -.78 & .47 \\ .37 & -.33 & -.87 \\ -.84 & -.52 & -.16 \end{bmatrix} \begin{bmatrix} 7.10 & 0 & 0 \\ 0 & 3.10 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} .30 & -.51 & -.81 \\ .76 & .64 & -.12 \\ .58 & -.58 & .58 \end{bmatrix}$$

a. ¿Cuál es el rango de A?

b. Utilice esta descomposición de A, sin realizar cálculos, para escribir una base para Col A y una base para Nul A. [*Sugerencia:* Primero escriba las columnas de V.]

16. Repita el ejercicio 15 para la siguiente DVS de una matriz A de 3×4:

$$A = \begin{bmatrix} -.86 & -.11 & -.50 \\ .31 & .68 & -.67 \\ .41 & -.73 & -.55 \end{bmatrix} \begin{bmatrix} 12.48 & 0 & 0 & 0 \\ 0 & 6.34 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} .66 & -.03 & -.35 & .66 \\ -.13 & -.90 & -.39 & -.13 \\ .65 & .08 & -.16 & -.73 \\ -.34 & .42 & -.84 & -.08 \end{bmatrix}$$

En los ejercicios 17 a 24, A es una matriz de $m \times n$ con una descomposición en valores singulares $A = U\Sigma V^T$, donde U es una matriz ortogonal de $m \times m$, Σ es una matriz "diagonal" de $m \times n$ con r entradas positivas y sin entradas negativas, y V es una matriz ortogonal de $n \times n$. Justifique sus respuestas.

17. Suponga que A es cuadrada e invertible. Encuentre una descomposición en valores singulares para A^{-1}.

18. Demuestre que si A es cuadrada, entonces $|\det A|$ es el producto de los valores singulares de A.

19. Demuestre que las columnas de V son vectores propios de A^TA, que las columnas de U son vectores propios de AA^T, y que las entradas diagonales de Σ son los valores singulares de A. [*Sugerencia:* Utilice la DVS para calcular A^TA y AA^T.]

20. Muestre que si A es una matriz de $n \times n$ y definida positiva, entonces una diagonalización ortogonal $A = PDP^T$ es una descomposición en valores singulares de A.

21. Muestre que si P es una matriz ortogonal de $m \times m$, entonces PA tiene los mismos valores singulares que A.

22. Justifique el enunciado del ejemplo 2 acerca de que el segundo valor singular de una matriz A es el máximo de $\|A\mathbf{x}\|$ cuando $\mathbf{x}$ varía sobre todos los vectores unitarios ortogonales a $\mathbf{v}_1$, siendo $\mathbf{v}_1$ un vector singular derecho correspondiente al primer valor singular de A. [*Sugerencia:* Utilice el teorema 7 de la sección 7.3.]

23. Si $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m]$ y $V = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$, muestre que

$$\boxed{A = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T}$$

24. Usando la notación del ejercicio 23, muestre que $A^T\mathbf{u}_j = \sigma_j\mathbf{v}_j$ para $1 \le j \le r = \text{rango } A$.

25. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Describa cómo encontrar una base $\mathcal{B}$ para $\mathbb{R}^n$ y una base $\mathcal{C}$ para $\mathbb{R}^m$ tales que la matriz para T relativa a $\mathcal{B}$ y $\mathcal{C}$ sea una matriz "diagonal" de $m \times n$.

[M] Calcule una DVS de cada matriz de los ejercicios 26 y 27. Escriba las entradas de la matriz final con hasta dos posiciones decimales. Utilice el método de los ejemplos 3 y 4.

26. $A = \begin{bmatrix} -18 & 13 & -4 & 4 \\ 2 & 19 & -4 & 12 \\ -14 & 11 & -12 & 8 \\ -2 & 21 & 4 & 8 \end{bmatrix}$

27. $A = \begin{bmatrix} 6 & -8 & -4 & 5 & -4 \\ 2 & 7 & -5 & -6 & 4 \\ 0 & -1 & -8 & 2 & 2 \\ -1 & -2 & 4 & 4 & -8 \end{bmatrix}$

28. **[M]** Encuentre los valores singulares de la matriz de 4×4 del ejercicio 9, sección 2.3, y calcule el número de condición σ_1/σ_4.

29. **[M]** Encuentre los valores singulares de la matriz de 5×5 del ejercicio 10, sección 2.3, y calcule el número de condición σ_1/σ_5.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Si $A = U\Sigma V^T$, donde Σ es de $m \times n$, entonces $A^T = (V^T)^T\Sigma^TU^T = V\Sigma^TU^T$. Ésta es una DVS para A^T porque V y U son matrices ortogonales y Σ^T es una matriz "diagonal" de $n \times m$. Como Σ y Σ^T tienen las mismas entradas diagonales diferentes de cero, A y A^T tienen los mismos valores singulares diferentes de cero. [*Nota:* Si A es de $2 \times n$, entonces AA^T es de solamente de 2×2, y sus valores propios pueden ser más fáciles de calcular (a mano) que los valores propios de A^TA.]

7.5 APLICACIONES AL PROCESAMIENTO DE IMÁGENES Y A LA ESTADÍSTICA

Las fotografías de satélite que aparecen en la introducción al capítulo proporcionan un ejemplo de datos multidimensionales o *multivariados* —información organizada de manera que cada dato del conjunto de datos se identifica mediante un punto (vector) en

$\mathbb{R}^n$. El principal objetivo de esta sección es explicar una técnica, llamada *análisis de componentes principales*, que se usa para analizar tales datos multivariados. Los cálculos ilustrarán el uso de la diagonalización ortogonal y de la descomposición en valores singulares.

El análisis de componentes principales puede aplicarse a cualesquiera datos que consistan en listas de mediciones efectuadas a una colección de objetos o individuos. Por ejemplo, considere un proceso químico que produce cierto material plástico. Para vigilar el proceso, se toman 300 muestras del material producido y cada muestra se somete a una serie de ocho pruebas, tales como punto de fusión, densidad, ductilidad, resistencia a la tensión, y otras. El informe de laboratorio para cada muestra es un vector en $\mathbb{R}^8$, y el conjunto de tales vectores forma una matriz de 8×300 llamada **matriz de observaciones**.

En términos simples, puede afirmarse que los datos de control del proceso son de dimensión 8. Los siguientes dos ejemplos describen datos que pueden visualizarse de manera gráfica.

EJEMPLO 1 Un ejemplo de datos bidimensionales está dado por un conjunto de pesos y estaturas de N estudiantes de licenciatura. Denote con $\mathbf{X}_j$ el **vector de observación** en $\mathbb{R}^2$ que enlista el peso y la estatura del j-ésimo estudiante. Si w denota el peso y h la estatura, entonces la matriz de observaciones tiene la forma

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_N \\ h_1 & h_2 & \cdots & h_N \end{bmatrix}$$
$$\begin{matrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{X}_1 & \mathbf{X}_2 & & \mathbf{X}_N \end{matrix}$$

El conjunto de vectores de observación puede visualizarse como un *diagrama de dispersión* bidimensional. Vea la figura 1.

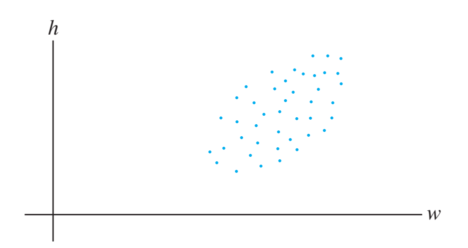


FIGURA 1 Diagrama de dispersión de vectores de observación $\mathbf{X}_1, \ldots, \mathbf{X}_N$.

EJEMPLO 2 Las primeras tres fotografías del valle Railroad en Nevada, Estados Unidos, mostradas en la introducción del capítulo, se pueden ver como *una* imagen de la región, con *tres componentes espectrales*, porque se hicieron mediciones simultáneas del lugar en tres longitudes de onda distintas. Cada fotografía proporciona información diferente sobre la misma área física. Por ejemplo, el primer píxel de la esquina superior izquierda de cada fotografía corresponde al mismo lugar en el suelo (de unos 30 por 30 metros.) A cada píxel le corresponde un vector de observación en $\mathbb{R}^3$ que enlista las intensidades de señal para ese píxel en las tres bandas espectrales.

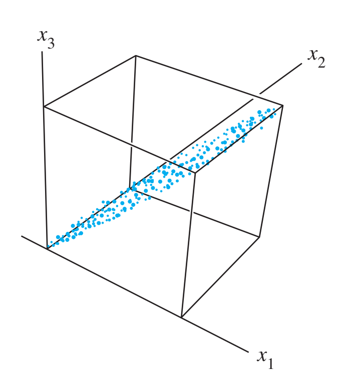


FIGURA 2 Diagrama de dispersión de datos espectrales para una imagen de satélite.

En forma típica, la imagen es de 2 000 × 2 000 píxeles, así que hay 4 millones de píxeles en la imagen. Los datos para la imagen forman una matriz con 3 filas y 4 millones de columnas (con las columnas acomodadas en cualquier orden conveniente). En este caso, el carácter "multidimensional" de los datos se refiere a las tres dimensiones *espectrales* más que a las dos dimensiones *espaciales* pertenecientes, de manera natural, a cualquier fotografía. Los datos pueden verse como un aglomerado de 4 millones de puntos en $\mathbb{R}^3$, quizá como en la figura 2.

Media y covarianza

En preparación para el análisis de componentes principales, sea $[\mathbf{X}_1 \ \cdots \ \mathbf{X}_N]$ una matriz de observaciones de $p \times N$, tal como se describió anteriormente. La **media muestral**, **M**, de los vectores de observación $\mathbf{X}_1, \ldots, \mathbf{X}_N$ está dada por

$$\mathbf{M} = \frac{1}{N}(\mathbf{X}_1 + \cdots + \mathbf{X}_N)$$

Para los datos de la figura 1, la media muestral es el punto ubicado en el "centro" del diagrama de dispersión. Para $k = 1, \ldots, N$, sea

$$\hat{\mathbf{X}}_k = \mathbf{X}_k - \mathbf{M}$$

Las columnas de la matriz de $p \times N$

$$B = [\,\hat{\mathbf{X}}_1 \quad \hat{\mathbf{X}}_2 \quad \cdots \quad \hat{\mathbf{X}}_N\,]$$

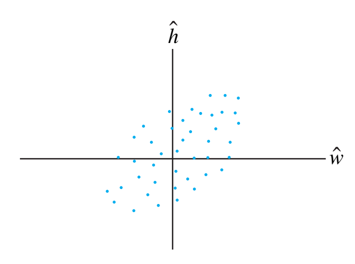


FIGURA 3 Datos de estatura y peso en forma de desviación media.

tienen una media muestral de cero, y se afirma que B está en **forma de desviación media**. Cuando la media muestral se resta a los datos de la figura 1, el diagrama de dispersión resultante tiene la forma que muestra la figura 3.

La **matriz de covarianza** (**muestral**) es la matriz S de $p \times p$ definida mediante

$$S = \frac{1}{N-1} B B^T$$

Dado que toda matriz de la forma BB^T es semidefinida positiva, también lo es S. (Vea el ejercicio 25 de la sección 7.2 con B y B^T intercambiadas.)

EJEMPLO 3 En un muestreo aleatorio de cierta población, se toman tres medidas de cada uno de cuatro individuos. Los vectores de observación son

$$\mathbf{X}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix}, \quad \mathbf{X}_3 = \begin{bmatrix} 7 \\ 8 \\ 1 \end{bmatrix}, \quad \mathbf{X}_4 = \begin{bmatrix} 8 \\ 4 \\ 5 \end{bmatrix}$$

Encuentre la media muestral y la matriz de covarianza.

Solución La media muestral es

$$\mathbf{M} = \frac{1}{4}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 1 \end{bmatrix} + \begin{bmatrix} 8 \\ 4 \\ 5 \end{bmatrix} \right) = \frac{1}{4}\begin{bmatrix} 20 \\ 16 \\ 20 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ 5 \end{bmatrix}$$

Al restar la media muestral a $\mathbf{X}_1, \ldots, \mathbf{X}_4$ se obtiene

$$\hat{\mathbf{X}}_1 = \begin{bmatrix} -4 \\ -2 \\ -4 \end{bmatrix}, \quad \hat{\mathbf{X}}_2 = \begin{bmatrix} -1 \\ -2 \\ 8 \end{bmatrix}, \quad \hat{\mathbf{X}}_3 = \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix}, \quad \hat{\mathbf{X}}_4 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

y

$$B = \begin{bmatrix} -4 & -1 & 2 & 3 \\ -2 & -2 & 4 & 0 \\ -4 & 8 & -4 & 0 \end{bmatrix}$$

La matriz de covarianza muestral es

$$S = \frac{1}{3} \begin{bmatrix} -4 & -1 & 2 & 3 \\ -2 & -2 & 4 & 0 \\ -4 & 8 & -4 & 0 \end{bmatrix} \begin{bmatrix} -4 & -2 & -4 \\ -1 & -2 & 8 \\ 2 & 4 & -4 \\ 3 & 0 & 0 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 30 & 18 & 0 \\ 18 & 24 & -24 \\ 0 & -24 & 96 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 0 \\ 6 & 8 & -8 \\ 0 & -8 & 32 \end{bmatrix}$$

Para analizar las entradas de $S = [s_{ij}]$, sea $\mathbf{X}$ tal que represente un vector que varía sobre el conjunto de vectores de observación, y denote las coordenadas de $\mathbf{X}$ con $x_1, \ldots, x_p$. Entonces x_1, por ejemplo, es un escalar que varía sobre el conjunto de las primeras coordenadas de $\mathbf{X}_1, \ldots, \mathbf{X}_N$. Para $j = 1, \ldots, p$, la entrada diagonal s_{jj} de S se llama **varianza** de x_j.

La varianza de x_j mide la dispersión de los valores de x_j. (Vea el ejercicio 13.) En el ejemplo 3, la varianza de x_1 es 10 y la de x_3 es 32. El que 32 sea más que 10 indica que el conjunto de las terceras entradas en los vectores de respuesta tiene una dispersión más amplia de valores que el conjunto de las primeras entradas.

La **varianza total** de los datos es la suma de las varianzas encontradas en la diagonal de S. En general, la suma de las entradas diagonales de una matriz cuadrada S se llama **traza** de la matriz, y se escribe $\text{tr}(S)$. Entonces

$$\{\text{varianza total}\} = \text{tr}(S)$$

La entrada s_{ij} de S para $i \neq j$ se llama **covarianza** de x_i y x_j. Observe que en el ejemplo 3, la covarianza entre x_1 y x_3 es 0 porque la entrada (1, 3) de S es 0. Los estadísticos afirman que x_1 y x_3 **no están correlacionadas**. El análisis de datos multivariados en $\mathbf{X}_1, \ldots, \mathbf{X}_N$ se simplifica mucho cuando la mayor parte de todas las variables $x_1, \ldots, x_p$ no están correlacionadas; esto es, cuando la matriz de covarianza de $\mathbf{X}_1, \ldots, \mathbf{X}_N$ es diagonal o casi diagonal.

Análisis de componentes principales

En aras de la simplicidad, suponga que la matriz $[\mathbf{X}_1 \ \cdots \ \mathbf{X}_N]$ ya está en forma de desviación media. El objetivo del análisis de componentes principales es encontrar una

matriz ortogonal de $p \times p$ $P = [\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_p]$ que determine un cambio de variable, $\mathbf{X} = P\mathbf{Y}$, o bien

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\,] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}$$

con la propiedad de que las nuevas variables $y_1, \ldots, y_p$ no están correlacionadas y están acomodadas en orden de varianza descendente.

El cambio ortogonal de variable $\mathbf{X} = P\mathbf{Y}$ significa que cada vector de observación $\mathbf{X}_k$ recibe un "nuevo nombre", $\mathbf{Y}_k$, tal que $\mathbf{X}_k = P\mathbf{Y}_k$. Observe que $\mathbf{Y}_k$ es el vector de coordenadas de $\mathbf{X}_k$ con respecto a las columnas de P, y que $\mathbf{Y}_k = P^{-1}\mathbf{X}_k = P^T\mathbf{X}_k$ para $k = 1, \ldots, N$.

No resulta difícil verificar que para cualquier P ortogonal, la matriz de covarianza de $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ es P^TSP (ejercicio 11). Entonces la matriz ortogonal P deseada es aquella que vuelve diagonal a P^TSP. Sea D una matriz diagonal con los valores propios $\lambda_1, \ldots, \lambda_p$ de S en la diagonal, acomodados de manera que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, y sea P una matriz ortogonal cuyas columnas son los vectores propios unitarios correspondientes $\mathbf{u}_1, \ldots, \mathbf{u}_p$. Entonces $S = PDP^T$ y $P^TSP = D$.

Los vectores propios unitarios $\mathbf{u}_1, \ldots, \mathbf{u}_p$ de la matriz de covarianza S se llaman **componentes principales** de los datos (en la matriz de observaciones). El **primer componente principal** es el vector propio correspondiente al mayor valor propio de S, el **segundo componente principal** es el vector propio correspondiente al segundo mayor valor propio, y así por el estilo.

El primer componente principal $\mathbf{u}_1$ determina la nueva variable y_1 de la siguiente forma. Sean $c_1, \ldots, c_p$ las entradas de $\mathbf{u}_1$. Como $\mathbf{u}_1{}^T$ es la primera fila de P^T, la ecuación $\mathbf{Y} = P^T\mathbf{X}$ muestra que

$$y_1 = \mathbf{u}_1^T\mathbf{X} = c_1x_1 + c_2x_2 + \cdots + c_px_p$$

Entonces y_1 es una combinación lineal de las variables originales $x_1, \ldots, x_p$, para la cual se usan las entradas del vector propio $\mathbf{u}_1$ como pesos. De manera similar, $\mathbf{u}_2$ determina la variable y_2, y así por el estilo.

EJEMPLO 4 Los datos iniciales para la imagen multiespectral del valle Railroad (ejemplo 2) consistían en 4 millones de vectores en $\mathbb{R}^3$. La matriz de covarianza asociada[1] es

$$S = \begin{bmatrix} 2382.78 & 2611.84 & 2136.20 \\ 2611.84 & 3106.47 & 2553.90 \\ 2136.20 & 2553(90 & 2650.71 \end{bmatrix}$$

Encuentre los componentes principales de los datos, y escriba la nueva variable determinada mediante los primeros componentes principales.

Solución Los valores propios de S y los componentes principales asociados (los vectores propios unitarios) son

[1]Los datos para formular el ejemplo 4 y los ejercicios 5 y 6 fueron proporcionados por Earth Satellite Corporation de Rockville, Maryland, Estados Unidos.

$$\lambda_1 = 7614.23 \qquad \lambda_2 = 427.63 \qquad \lambda_3 = 98.10$$

$$\mathbf{u}_1 = \begin{bmatrix} .5417 \\ .6295 \\ .5570 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} -.4894 \\ -.3026 \\ .8179 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} .6834 \\ -.7157 \\ .1441 \end{bmatrix}$$

Si, en aras de la simplicidad, se usan dos posiciones decimales, la variable para el primer componente principal es

$$y_1 = .54x_1 + .63x_2 + .56x_3$$

Esta ecuación se usó para crear la fotografía (d) que aparece en la introducción al capítulo. Las variables x_1, x_2, x_3 son las intensidades de señal presentes en las tres bandas espectrales. Los valores de x_1, convertidos a una escala de grises entre el negro y el blanco, produjeron la fotografía (a). De manera similar, los valores de x_2 y x_3 produjeron las fotografías (b) y (c), respectivamente. En cada píxel de la fotografía (d), se calculó el valor de la escala de grises a partir de y_1, una combinación lineal ponderada de x_1, x_2, x_3. En este sentido, la fotografía (d) "despliega" el primer componente principal de los datos.

En el ejemplo 4, la matriz de covarianza para los datos transformados, usando las variables y_1, y_2, y_3, es

$$D = \begin{bmatrix} 7614.23 & 0 & 0 \\ 0 & 427.63 & 0 \\ 0 & 0 & 98.10 \end{bmatrix}$$

Aunque D es, desde luego, más simple que la matriz de covarianza original S, todavía no resulta evidente la ventaja de estructurar las nuevas variables. Sin embargo, las varianzas de las variables y_1, y_2, y_3 aparecen en la diagonal de D y, evidentemente la primera varianza de D es mucho mayor que las otras dos. Como se verá, esto permite ver los datos esencialmente como unidimensionales en vez de tridimensionales.

Reducción de la dimensión de datos multivariados

El análisis de componentes principales puede resultar valioso para aplicaciones en que la mayor parte de la variación, o intervalo dinámico, de los datos se debe a variaciones de *sólo unas cuantas* de las nuevas variables, $y_1, \ldots, y_p$.

Puede demostrarse que un cambio ortogonal de variables, $\mathbf{X} = P\mathbf{Y}$, no cambia la varianza total de los datos. (En términos generales, esto es cierto porque la multiplicación izquierda por P no altera las longitudes de los vectores ni los ángulos entre ellos. Vea el ejercicio 12.) Esto significa que si $S = PDP^T$, entonces

$$\left\{\begin{matrix}\text{varianza total} \\ \text{de } x_1, \ldots, x_p\end{matrix}\right\} = \left\{\begin{matrix}\text{varianza total} \\ \text{de } y_1, \ldots, y_p\end{matrix}\right\} = \operatorname{tr}(D) = \lambda_1 + \cdots + \lambda_p$$

La varianza de y_j es λ_j, y el cociente $\lambda_j/\operatorname{tr}(S)$ mide la fracción de la varianza total que se "explica" o "captura" mediante y_j.

EJEMPLO 5 Encuentre los diversos porcentajes de varianza para los datos multiespectrales del valle Railroad que se desplegaron en las fotografías de los componentes principales, (d), (e) y (f), mostrados en la introducción del capítulo.

Solución La varianza total de los datos es

$$\text{tr}(D) = 7614.23 + 427.63 + 98.10 = 8139.96$$

[Verifique si este número también es igual a tr(S).] Los porcentajes de la varianza total explicados mediante los componentes principales son

Primer componente	Segundo componente	Tercer componente
$\frac{7614.23}{8139.96} = 93.5\%$	$\frac{427.63}{8139.96} = 5.3\%$	$\frac{98.10}{8139.96} = 1.2\%$

En cierto sentido, el 93.5% de la información recopilada por medio de Landsat para la región del valle Railroad se exhibe en la fotografía (d), con el 5.3% en (e) y sólo el 1.2% restante en (f).

Los cálculos del ejemplo 5 muestran que los datos prácticamente no tienen varianza en la tercera coordenada (nueva). Todos los valores de y_3 son cercanos a cero. Geométricamente, los puntos de datos están cercanos al plano $y_3 = 0$, y es posible determinar, con relativa precisión, dónde se ubican conociendo solamente los valores de y_1 y y_2. De hecho, y_2 también tiene una varianza relativamente pequeña, lo cual significa que los puntos están aproximadamente a lo largo de una línea, y que los datos son esencialmente unidimensionales. Vea la figura 2, en la cual los datos tienen la apariencia de un palillo de paleta.

Caracterizaciones de variables de componentes principales

Si $y_1, \dots, y_p$ surge del análisis de componentes principales de una matriz de observaciones de $p \times N$, entonces la varianza de y_1 es tan grande como es posible en el siguiente sentido: si $\mathbf{u}$ es cualquier vector unitario y si $y = \mathbf{u}^T\mathbf{X}$, entonces la varianza de los valores de y cuando $\mathbf{X}$ varía sobre los datos originales $\mathbf{X}_1, \dots, \mathbf{X}_N$ resulta ser $\mathbf{u}^T S\mathbf{u}$. De acuerdo con el teorema 8 de la sección 7.3, el valor máximo de $\mathbf{u}^T S\mathbf{u}$, sobre todos los vectores unitarios $\mathbf{u}$, es el mayor valor propio λ_1 de S, y se alcanza esta varianza cuando $\mathbf{u}$ es el vector propio correspondiente $\mathbf{u}_1$. De igual forma, el teorema 8 muestra que y_2 tiene la máxima varianza posible entre todas las variables $y = \mathbf{u}^T\mathbf{X}$ que *no* están *correlacionadas* con y_1. Asimismo, y_3 tiene la máxima varianza posible entre todas las variables no correlacionadas tanto con y_1 como con y_2, y así sucesivamente.

NOTA NUMÉRICA

La descomposición en valores singulares es la herramienta fundamental para realizar el análisis de componentes principales en aplicaciones prácticas. Si B es una matriz de observaciones de $p \times N$ en la forma de desviación media, y si $A = \left(1/\sqrt{N-1}\right) B^T$, entonces A^TA es la matriz de covarianza S. Los cuadrados de los valores singulares de A son los p valores propios de S, y los vectores singulares derechos de A son los componentes principales de los datos.

Como se mencionó en la sección 7.4, el cálculo iterativo de la DVS para A es más rápido y preciso que una descomposición en valores propios de S. Esto es particularmente cierto, por ejemplo, en el procesamiento de imágenes hiperespectral (con $p = 224$) mencionado en la introducción al capítulo. El análisis de componentes principales se realiza en segundos en estaciones de trabajo especializadas.

Lectura adicional

Lillesand, Thomas M. y Ralph W. Kiefer, *Remote Sensing and Image Interpretation*, 4a. ed. (Nueva York: John Wiley, 2000).

PROBLEMAS DE PRÁCTICA

La tabla siguiente enlista los pesos y estaturas de cinco muchachos:

Muchacho	#1	#2	#3	#4	#5
Peso (lb)	120	125	125	135	145
Estatura (pulg)	61	60	64	68	72

1. Encuentre la matriz de covarianza para los datos.
2. Efectúe un análisis de componentes principales de los datos para encontrar un único *índice de tamaño* que explique la mayor parte de la variación de los datos.

7.5 EJERCICIOS

En los ejercicios 1 y 2, convierta la matriz de observaciones a la forma de desviación media y estructure la matriz de covarianza muestral.

1. $\begin{bmatrix} 19 & 22 & 6 & 3 & 2 & 20 \\ 12 & 6 & 9 & 15 & 13 & 5 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 5 & 2 & 6 & 7 & 3 \\ 3 & 11 & 6 & 8 & 15 & 11 \end{bmatrix}$

3. Encuentre los componentes principales de los datos del ejercicio 1.

4. Encuentre los componentes principales de los datos del ejercicio 2.

5. **[M]** Se estructuró una imagen Landsat con tres componentes espectrales de la base de la Fuerza Aérea estadounidense Homestead en Florida (luego de que el huracán Andrew azotara esta base en 1992). A continuación se muestra la matriz de covarianza de los datos. Encuentre el primer componente principal de los datos, y calcule el porcentaje de la varianza total contenida en este componente.

$$S = \begin{bmatrix} 164.12 & 32.73 & 81.04 \\ 32.73 & 539.44 & 249.13 \\ 81.04 & 249.13 & 189.11 \end{bmatrix}$$

6. **[M]** La siguiente matriz de covarianza se obtuvo de una imagen Landsat del río Columbia en Washington, EUA, utilizando datos de tres bandas espectrales. Sean x_1, x_2, x_3 los componentes espectrales de cada píxel de la imagen. Encuentre una nueva variable de la forma $y_1 = c_1x_1 + c_2x_2 + c_3x_3$ que tenga la máxima varianza posible, sujeta a la restricción de que $c_1^2 + c_2^2 + c_3^2 = 1$. ¿Qué porcentaje de la varianza total de los datos se explica mediante y_1?

$$S = \begin{bmatrix} 29.64 & 18.38 & 5.00 \\ 18.38 & 20.82 & 14.06 \\ 5.00 & 14.06 & 29.21 \end{bmatrix}$$

7. Denote con x_1, x_2 las variables para los datos bidimensionales del ejercicio 1. Encuentre una nueva variable y_1 de la forma $y_1 = c_1x_1 + c_2x_2$, con $c_1^2 + c_2^2 = 1$, tal que y_1 tenga la máxima varianza posible sobre los datos dados. ¿Qué porcentaje de la varianza en los datos se explica mediante y_1?

8. Repita el ejercicio 7 con los datos del ejercicio 2.

9. Suponga que se aplican tres pruebas a una muestra de estudiantes de licenciatura. Sean $\mathbf{X}_1, \ldots, \mathbf{X}_N$ los vectores de observación en $\mathbb{R}^3$ que enlistan las tres puntuaciones por estudiante, y para $j = 1, 2, 3$, denote con x_j la puntuación de un estudiante en la j-ésima prueba. Suponga que la matriz de covarianza de los datos es

$$S = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}$$

Sea y un "índice" del desempeño estudiantil, con $y = c_1x_1 + c_2x_2 + c_3x_3$ y $c_1^2 + c_2^2 + c_3^2 = 1$. Elija c_1, c_2, c_3 de manera que la varianza de y sobre el conjunto de datos sea lo más grande posible. [*Pista:* Los valores propios de la matriz de covarianza muestral son $\lambda = 3, 6,$ y 9.]

10. **[M]** Repita el ejercicio 9 con $S = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 11 & 4 \\ 2 & 4 & 5 \end{bmatrix}$.

11. Dados los datos multivariados $\mathbf{X}_1, \ldots, \mathbf{X}_N$ (en $\mathbb{R}^p$) en forma de desviación media, sea P una matriz de $p \times p$, y defínase $\mathbf{Y}_k = P^T\mathbf{X}_k$ para $k = 1, \ldots, N$.

a. Muestre que $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ están en forma de desviación media. [*Sugerencia:* Sea $\mathbf{w}$ el vector en $\mathbb{R}^N$ con un 1 en cada entrada. Entonces $[\mathbf{X}_1 \ \cdots \ \mathbf{X}_N]\mathbf{w} = \mathbf{0}$ (el vector cero en $\mathbb{R}^p$).]

b. Muestre que si la matriz de covarianza de $\mathbf{X}_1, \ldots, \mathbf{X}_N$ es S, entonces la matriz de covarianza de $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ es P^TSP.

12. Denote por medio de $\mathbf{X}$ un vector que varía sobre las columnas de una matriz de observación de $p \times N$, y sea P una matriz ortogonal de $p \times p$. Muestre que el cambio de variable $\mathbf{X} = P\mathbf{Y}$ no cambia la varianza total de los datos. [*Sugerencia:* Según el ejercicio 11, es suficiente mostrar que tr $(P^TSP) =$ tr (S). Use una propiedad de la traza mencionada en el ejercicio 25 de la sección 5.4.]

13. La matriz de covarianza muestral es una generalización de una fórmula para la varianza muestral de N mediciones muestrales, por ejemplo, $t_1, \ldots, t_N$. Si m es el promedio de $t_1, \ldots, t_N$, entonces la *varianza muestral* está dada por

$$\frac{1}{N-1}\sum_{k=1}^{n}(t_k - m)^2 \tag{1}$$

Muestre cómo la matriz de covarianza muestral, S, definida antes del ejemplo 3, puede escribirse en una forma similar a (1). [*Sugerencia:* Utilice la multiplicación de matrices partidas para escribir S como $1/(N-1)$ veces la suma de N matrices de tamaño $p \times p$. Para $1 \le k \le N$, escriba $\mathbf{X}_k - \mathbf{M}$ en lugar de $\hat{\mathbf{X}}_k$.]

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Primero, acomode los datos en forma de desviación media. Resulta fácil advertir que el vector de la media muestral es $\mathbf{M} = \begin{bmatrix} 130 \\ 65 \end{bmatrix}$. Reste $\mathbf{M}$ a los vectores de observación (las columnas de la tabla) y obtenga

$$B = \begin{bmatrix} -10 & -5 & -5 & 5 & 15 \\ -4 & -5 & -1 & 3 & 7 \end{bmatrix}$$

Entonces la matriz de covarianza muestral es

$$S = \frac{1}{5-1}\begin{bmatrix} -10 & -5 & -5 & 5 & 15 \\ -4 & -5 & -1 & 3 & 7 \end{bmatrix}\begin{bmatrix} -10 & -4 \\ -5 & -5 \\ -5 & -1 \\ 5 & 3 \\ 15 & 7 \end{bmatrix}$$

$$= \frac{1}{4}\begin{bmatrix} 400 & 190 \\ 190 & 100 \end{bmatrix} = \begin{bmatrix} 100.0 & 47.5 \\ 47.5 & 25.0 \end{bmatrix}$$

2. Los valores propios de S son (hasta dos posiciones decimales)

$$\lambda_1 = 123.02 \quad \text{y} \quad \lambda_2 = 1.98$$

El vector propio unitario correspondiente a λ_1 es $\mathbf{u} = \begin{bmatrix} .900 \\ .436 \end{bmatrix}$. (Como S es de 2×2, los cálculos pueden hacerse a mano si no se tiene un programa de matrices.) Para el *índice de tamaño*, sea

$$y = .900\hat{w} + .436\hat{h}$$

donde $\hat{w}$ y $\hat{h}$ son peso y estatura, respectivamente, en forma de desviación media. La varianza de este índice sobre el conjunto de los datos es de 123.02. Debido a que la varianza total es $\text{tr}(S) = 100 + 25 = 125$, el índice de tamaño cubre prácticamente toda la varianza de los datos (98.4%).

Los datos originales para el problema de práctica 1 y la línea determinada por el primer componente principal **u** se muestran en la figura 4. (En forma de vector paramétrico, la línea es $\mathbf{x} = \mathbf{M} + t\mathbf{u}$.) Puede mostrarse que la línea es la mejor aproximación a los datos, en el sentido de que la suma de los cuadrados de las distancias que son *ortogonales* a la línea se minimiza. De hecho, el análisis de componentes principales equivale a lo que se llama *regresión ortogonal*, pero ésa es una historia que se dejará para otra ocasión.

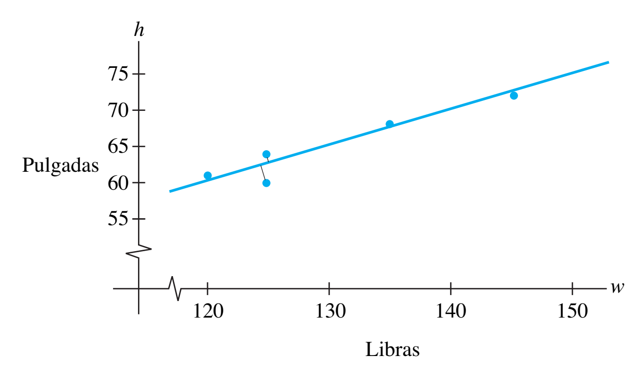


FIGURA 4 Línea de regresión ortogonal determinada por el primer componente principal de los datos.

CAPÍTULO 7 EJERCICIOS SUPLEMENTARIOS

1. Señale cada enunciado como verdadero o falso. Justifique sus respuestas. En cada inciso, A representa una matriz de $n \times n$.

a. Si A es diagonalizable ortogonalmente, entonces A es simétrica.

b. Si A es una matriz ortogonal, entonces A es simétrica.

c. Si A es una matriz ortogonal, entonces $\|A\mathbf{x}\| = \|\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

d. Los ejes principales de una forma cuadrática $\mathbf{x}^TA\mathbf{x}$ pueden ser las columnas de cualquier matriz P que diagonalice A.

e. Si P es una matriz de $n \times n$ con columnas ortogonales, entonces $P^T = P^{-1}$.

f. Si todo coeficiente de una forma cuadrática es positivo, entonces la forma cuadrática es definida positiva.

g. Si $\mathbf{x}^TA\mathbf{x} > 0$ para alguna $\mathbf{x}$, entonces la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ es definida positiva.

h. Por medio de un cambio de variable apropiado, cualquier forma cuadrática puede transformarse en una forma cuadrática sin términos de producto cruzado.

i. El mayor valor para una forma cuadrática $\mathbf{x}^TA\mathbf{x}$, para $\|\mathbf{x}\| = 1$, es la mayor entrada en la diagonal de A.

j. El valor máximo de una forma cuadrática $\mathbf{x}^TA\mathbf{x}$ definida positiva es el mayor valor propio de A.

k. Una forma cuadrática definida positiva puede transformarse en una forma definida negativa por medio de un cambio de variable apropiado $\mathbf{x} = P\mathbf{u}$, para alguna matriz ortogonal P.

l. Una forma cuadrática indefinida es una cuyos valores propios no están definidos.

m. Si P es una matriz ortogonal de $n \times n$, entonces el cambio de variable $\mathbf{x} = P\mathbf{u}$ transforma a $\mathbf{x}^TA\mathbf{x}$ en una forma cuadrática cuya matriz es $P^{-1}AP$.

n. Si U es de $m \times n$ con columnas ortogonales, entonces $UU^T\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre Col U.

o. Si B es de $m \times n$ y $\mathbf{x}$ es un vector unitario en $\mathbb{R}^n$, entonces $\|B\mathbf{x}\| \leq \sigma_1$, donde σ_1 es el primer valor singular de B.

p. Una descomposición en valores singulares de una matriz B de $m \times n$ puede escribirse como $B = P\Sigma Q$, donde P es una matriz ortogonal de $m \times m$, Q es una matriz ortogonal de $n \times n$, y Σ es una matriz "diagonal" de $m \times n$.

q. Si A es de $n \times n$, entonces A y A^TA tienen los mismos valores singulares.

2. Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ una base ortonormal para $\mathbb{R}^n$, y sean $\lambda_1, \ldots, \lambda_n$ cualesquiera escalares reales. Defina

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T$$

a. Muestre que A es simétrica.

b. Muestre que $\lambda_1, \ldots, \lambda_n$ son los valores propios de A.

3. Sea A una matriz simétrica de $n \times n$ con rango r. Explique por qué la descomposición espectral de A representa A como la suma de r matrices de rango 1.

4. Sea A una matriz simétrica de $n \times n$.

a. Muestre que $(\text{Col } A)^{\perp} = \text{Nul } A$. [*Sugerencia*: Vea la sección 6.1.]

b. Muestre que toda $\mathbf{y}$ en $\mathbb{R}^n$ puede escribirse en la forma $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, con $\hat{\mathbf{y}}$ en Col A y $\mathbf{z}$ en Nul A.

5. Muestre que si $\mathbf{v}$ es un vector propio de una matriz A de $n \times n$ y $\mathbf{v}$ corresponde a un valor propio de A diferente de cero, entonces $\mathbf{v}$ está en Col A. [*Sugerencia:* Utilice la definición de vector propio.]

6. Sea A una matriz simétrica de $n \times n$. Utilice el ejercicio 5 y una base de vectores propios para $\mathbb{R}^n$ para dar una segunda demostración de la descomposición en el ejercicio 4(b).

7. Demuestre que una matriz A de $n \times n$ es definida positiva si, y sólo si, A admite una *factorización Cholesky*, es decir, $A = R^TR$ para alguna matriz triangular superior invertible R cuyas entradas diagonales son todas positivas. [*Sugerencia:* Utilice una factorización QR y el ejercicio 26 de la sección 7.2.]

8. Utilice el ejercicio 7 para demostrar que si A es definida positiva, entonces tiene una factorización LU, $A = LU$, donde U tiene pivotes positivos en su diagonal. (Lo recíproco también es cierto.)

Si A es de $m \times n$, entonces la matriz $G = A^TA$ se denomina *matriz Gram de A*. En este caso, las entradas de G son los productos interiores de las columnas de A.

9. Muestre que la matriz Gram de cualquier matriz A es semidefinida positiva, con el mismo rango que A. (Vea los ejercicios de la sección 6.5.)

10. Muestre que si una matriz G de $n \times n$ es semidefinida positiva y tiene rango r, entonces G es la matriz Gram de alguna matriz A de $r \times n$. Esto se llama *factorización reveladora de rango* de G. [*Sugerencia:* Considere la descomposición espectral de G, y primero escriba G como BB^T para una matriz B de $n \times r$.]

11. Demuestre que cualquier matriz A de $n \times n$ admite una *descomposición polar* de la forma $A = PQ$, donde P es una matriz semidefinida positiva de $n \times n$ con el mismo rango que A y Q es una matriz ortogonal de $n \times n$. [*Sugerencia:* Utilice una descomposición en valores singulares, $A = U\Sigma V^T$, y observe que $A = (U\Sigma U^T)(UV^T)$.] Esta descomposición se usa, por ejemplo, en ingeniería mecánica para modelar la deformación de un material. La matriz P describe el estiramiento o compresión del material (en las direcciones de los vectores propios de P), y Q describe la rotación del material en el espacio.

Los ejercicios 12 a 14 se refieren a una matriz A de $m \times n$ con una descomposición en valores singulares reducida, $A = U_rDV_r^T$, y con la seudoinversa $A^+ = V_rD^{-1}U_r^T$.

12. Verifique las propiedades de A^+:

a. Para toda $\mathbf{y}$ en $\mathbb{R}^m$, $AA^+\mathbf{y}$ es la proyección ortogonal de $\mathbf{y}$ sobre Col A.

b. Para toda $\mathbf{x}$ en $\mathbb{R}^n$, $A^+A\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre Fil A.

c. $AA^+A = A$ y $A^+AA^+ = A^+$.

13. Suponga que la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, y sea $\mathbf{x}^+ = A^+\mathbf{b}$. De acuerdo con el ejercicio 23 de la sección 6.3, existe exactamente un vector $\mathbf{p}$ en Fil A tal que $A\mathbf{p} = \mathbf{b}$. Los siguientes pasos demuestran que $\mathbf{x}^+ = \mathbf{p}$ y que $\mathbf{x}^+$ es la *solución de longitud mínima* de $A\mathbf{x} = \mathbf{b}$.

a. Muestre que $\mathbf{x}^+$ está en Fil A. [*Sugerencia:* Escriba $\mathbf{b}$ como $A\mathbf{x}$ para alguna $\mathbf{x}$, y utilice el ejercicio 12.]

b. Muestre que $\mathbf{x}^+$ es una solución de $A\mathbf{x} = \mathbf{b}$.

c. Muestre que si $\mathbf{u}$ es cualquier solución de $A\mathbf{x} = \mathbf{b}$, entonces $\|\mathbf{x}^+\| \leq \|\mathbf{u}\|$, con igualdad sólo si $\mathbf{u} = \mathbf{x}^+$.

14. Dada cualquier $\mathbf{b}$ en $\mathbb{R}^m$, adapte el ejercicio 13 para mostrar que $A^+\mathbf{b}$ es la *solución por mínimos cuadrados de longitud mínima*. [*Sugerencia:* Considere la ecuación $A\mathbf{x} = \hat{\mathbf{b}}$, donde $\hat{\mathbf{b}}$ es la proyección ortogonal de $\mathbf{b}$ sobre Col A.]

[M] En los ejercicios 15 y 16, Construya la seudoinversa de A. Comience por utilizar un programa de matrices para producir la DVS de A, o, si no tiene un programa, comience con una diagonalización ortogonal de A^TA. Utilice la seudoinversa para resolver $A\mathbf{x} = \mathbf{b}$, para $\mathbf{b} = (6, -1, -4, 6)$, y sea $\hat{\mathbf{x}}$ la solución. Efectúe un cálculo para verificar que $\hat{\mathbf{x}}$ está en Fil A. Encuentre un vector $\mathbf{u}$ diferente de cero en Nul A, y verifique si $\|\hat{\mathbf{x}}\| < \|\hat{\mathbf{x}} + \mathbf{u}\|$, lo que debe ser cierto según el ejercicio 13(c).

15. $A = \begin{bmatrix} -3 & -3 & -6 & 6 & 1 \\ -1 & -1 & -1 & 1 & -2 \\ 0 & 0 & -1 & 1 & -1 \\ 0 & 0 & -1 & 1 & -1 \end{bmatrix}$

16. $A = \begin{bmatrix} 4 & 0 & -1 & -2 & 0 \\ -5 & 0 & 3 & 5 & 0 \\ 2 & 0 & -1 & -2 & 0 \\ 6 & 0 & -3 & -6 & 0 \end{bmatrix}$

APÉNDICE A

Unicidad de la forma escalonada reducida

TEOREMA **Unicidad de la forma escalonada reducida**

Toda matriz A de $m \times n$ es equivalente por filas a una única matriz U escalonada reducida.

DEMOSTRACIÓN La demostración aplica la idea presentada en la sección 4.3 acerca de que las columnas de matrices equivalentes por filas tienen exactamente las mismas relaciones de dependencia lineal.

El algoritmo de reducción por filas muestra que existe al menos una matriz U semejante. Suponga que A es equivalente por filas a las matrices U y V de forma escalonada reducida. La entrada diferente de cero ubicada más hacia la izquierda en una fila de U es un "1 principal". Llame a la ubicación de este 1 principal una posición pivote, y a la columna que lo contiene columna pivote. (Esta definición emplea solamente la naturaleza escalonada de U y V, y no supone la unicidad de la forma escalonada reducida.)

Las columnas pivote de U y V son precisamente las columnas diferentes de cero que *no* son linealmente dependientes de las columnas situadas a su izquierda. (Esta condición la satisface automáticamente una *primera* columna si es diferente de cero.) Como U y V son equivalentes por filas (donde ambas son equivalentes por filas a A), sus columnas tienen las mismas relaciones de dependencia. Por lo tanto, las columnas pivote de U y V están en las mismas posiciones. Si hay r columnas de este tipo, entonces, como U y V están en forma escalonada reducida, sus columnas pivote son las primeras r columnas de la matriz identidad $m \times m$. De ahí que las *columnas pivote correspondientes de U y V sean iguales*.

Por último, considere cualquier columna no pivote de U, por ejemplo la columna j. Esta columna es cero o bien una combinación lineal de las columnas pivote localizadas a su izquierda (porque esas columnas pivote son una base para el espacio generado por las columnas situadas a la izquierda de la columna j). Cualquiera de estos casos puede expresarse al escribir $U\mathbf{x} = \mathbf{0}$ para algún $\mathbf{x}$ cuya j-ésima entrada sea 1. Entonces también $V\mathbf{x} = \mathbf{0}$, lo cual postula que la columna j de V es cero o bien la *misma* combinación lineal de las columnas pivote de V ubicadas a *su* izquierda. Como las correspondientes columnas pivote de U y V son iguales, las columnas j de U y V también son iguales. Esto es válido para cualquier columna no pivote, así que $V = U$, lo cual demuestra que U es única. ■

APÉNDICE B

Números complejos

Un **número complejo** es un número escrito en la forma

$$z = a + bi$$

donde a y b son números reales e i es el símbolo formal que satisface la relación $i^2 = -1$. El número a es la **parte real** de z, denotada mediante Re z, y b es la **parte imaginaria** de z, denotada con Im z. Se considera que dos números complejos son iguales si, y sólo si, sus partes real e imaginaria son iguales. Por ejemplo, si $z = 5 + (-2)i$, entonces Re $z = 5$ e Im $z = -2$. Por simplicidad, se escribe $z = 5 - 2i$.

Se considera que un número real a es un tipo especial de número complejo, al identificar a con $a + 0i$. Más aún, las operaciones aritméticas con números reales pueden ampliarse al conjunto de los números complejos.

El **sistema de números complejos**, denotado mediante $\mathbb{C}$, es el conjunto de todos los números complejos, junto con las siguientes operaciones para la suma y la multiplicación:

$$(a + bi) + (c + di) = (a + c) + (b + d)i \tag{1}$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i \tag{2}$$

Estas reglas se reducen a la suma y la multiplicación comunes de números reales cuando b y d son cero en (1) y (2). Se comprueba fácilmente que las reglas comunes de aritmética para $\mathbb{R}$ son válidas también para $\mathbb{C}$. Por esta razón, la multiplicación se calcula generalmente por medio de una ampliación algebraica, como en el siguiente ejemplo.

EJEMPLO 1

$$\begin{aligned}(5 - 2i)(3 + 4i) &= 15 + 20i - 6i - 8i^2 \\ &= 15 + 14i - 8(-1) \\ &= 23 + 14i\end{aligned}$$

Esto es, cada término de $5 - 2i$ se multiplica por cada término de $3 + 4i$, se usa $i^2 = -1$, y se escribe el resultado en la forma $a + bi$.

La resta de los números complejos z_1 y z_2 se define como

$$z_1 - z_2 = z_1 + (-1)z_2$$

En particular, se escribe $-z$ en lugar de $(-1)z$.

El **conjugado** de $z = a + bi$ es un número complejo $\overline{z}$ (lea "zeta testado"), definido mediante

$$\overline{z} = a - bi$$

Se obtiene $\overline{z}$ de z al cambiar el signo de la parte imaginaria.

EJEMPLO 2 El conjugado de $-3 + 4i$ es $-3 - 4i$; se escribe $\overline{-3 + 4i} = -3 - 4i$.

Observe que si $z = a + bi$, entonces

$$z\overline{z} = (a + bi)(a - bi) = a^2 - abi + bai - b^2i^2 = a^2 + b^2 \tag{3}$$

Como $z\overline{z}$ es real y no negativo, tiene raíz cuadrada. El **valor absoluto** (o **módulo**) de z es el número real $|z|$ definido mediante

$$|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$$

Si z es un número real, entonces $z = a + 0i$, y $|z| = \sqrt{a^2}$, que es igual al valor absoluto ordinario de a.

A continuación se enlistan algunas útiles propiedades de los conjugados y de los valores absolutos; w y z denotan números complejos.

1. $\overline{z} = z$ si, y sólo si, z es un número real.
2. $\overline{w + z} = \overline{w} + \overline{z}$.
3. $\overline{wz} = \overline{w}\,\overline{z}$; en particular, $\overline{rz} = r\overline{z}$ si r es un número real.
4. $z\overline{z} = |z|^2 \geq 0$.
5. $|wz| = |w||z|$.
6. $|w + z| \leq |w| + |z|$.

Si $z \neq 0$, entonces $|z| > 0$ y z tienen un inverso multiplicativo, denotado mediante $1/z$ o z^{-1} y dado por

$$\frac{1}{z} = z^{-1} = \frac{\overline{z}}{|z|^2}$$

Por supuesto, un cociente w/z significa simplemente $w \cdot (1/z)$.

EJEMPLO 3 Sean $w = 3 + 4i$ y $z = 5 - 2i$. Calcule $z\bar{z}$, $|z|$, y w/z.

Solución De (3),

$$z\bar{z} = 5^2 + (-2)^2 = 25 + 4 = 29$$

Para el valor absoluto, $|z| = \sqrt{z\bar{z}} = \sqrt{29}$. Para calcular w/z, primero multiplique tanto el numerador como el denominador por $\bar{z}$, el conjugado del denominador. De acuerdo con la ecuación (3), esto elimina la i del denominador:

$$\begin{aligned}\frac{w}{z} &= \frac{3+4i}{5-2i} \\ &= \frac{3+4i}{5-2i}\cdot\frac{5+2i}{5+2i} \\ &= \frac{15+6i+20i-8}{5^2+(-2)^2} \\ &= \frac{7+26i}{29} \\ &= \frac{7}{29}+\frac{26}{29}i\end{aligned}$$

Interpretación geométrica

Cada número complejo $z = a + bi$ corresponde a un punto (a, b) en el plano $\mathbb{R}^2$, como en la figura 1. El eje horizontal se llama **eje real** porque los puntos $(a, 0)$ representados en él corresponden a los números reales. El eje vertical es el **eje imaginario** porque los puntos $(0, b)$ representados en él corresponden a los **números imaginarios puros** de la forma $0 + bi$, o simplemente bi. El conjugado de z es la imagen de z reflejada en el eje real. El valor absoluto de z es la distancia desde (a, b) hasta el origen.

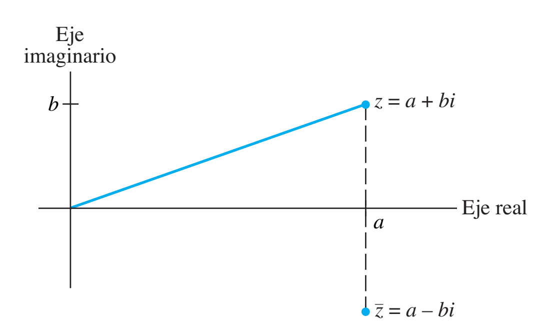


FIGURA 1 El conjugado complejo es una imagen reflejada.

La suma de los números complejos $z = a + bi$ y $w = c + di$ corresponde al vector suma de (a, b) y (c, d) en $\mathbb{R}^2$, como en la figura 2.

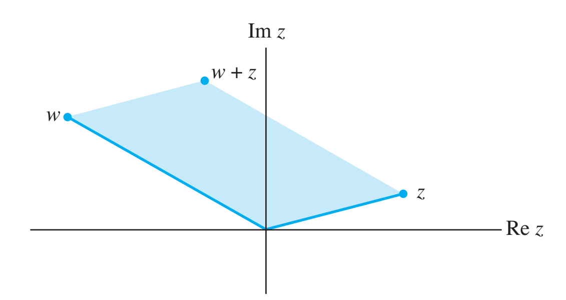


FIGURA 2 Suma de números complejos.

Para ofrecer una representación gráfica de la multiplicación de números complejos, se utilizan **coordenadas polares** en $\mathbb{R}^2$. Dado un número complejo diferente de cero $z = a + bi$, sea φ el ángulo entre el eje real positivo y el punto (a, b), como en la figura 3 donde $-\pi < \varphi \leq \pi$. El ángulo φ es el **argumento** de z; se escribe $\varphi = \arg z$. Por trigonometría,

$$a = |z| \cos \varphi, \qquad b = |z| \operatorname{sen} \varphi$$

y así

$$z = a + bi = |z|(\cos \varphi + i \operatorname{sen} \varphi)$$

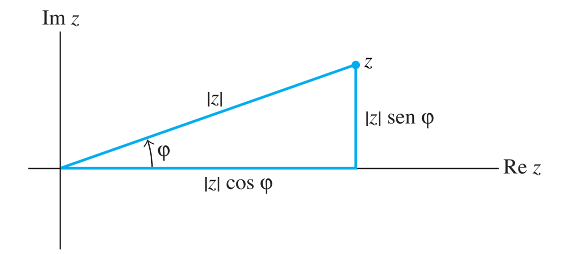


FIGURA 3 Coordenadas polares de z.

Si w es otro número complejo diferente de cero, por ejemplo,

$$w = |w| (\cos \vartheta + i \operatorname{sen} \vartheta)$$

entonces, utilizando identidades trigonométricas estándar para el seno y el coseno de la suma de dos ángulos, puede verificarse que

$$wz = |w|\,|z|\,[\cos(\vartheta + \varphi) + i \operatorname{sen}(\vartheta + \varphi)] \tag{4}$$

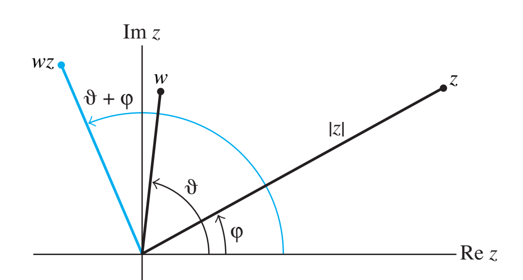
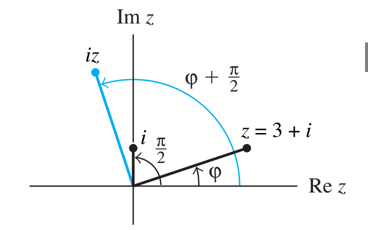


FIGURA 4 Multiplicación con coordenadas polares.

Vea la figura 4. Se puede escribir una fórmula similar para cocientes en la forma polar. Las fórmulas para productos y cocientes se pueden establecer en palabras de la manera siguiente.

> El producto de dos números complejos diferentes de cero está dado en forma polar por el producto de sus valores absolutos y la suma de sus argumentos. El cociente de dos números complejos diferentes de cero está dado por el cociente de sus valores absolutos y la diferencia de sus argumentos.



Multiplicación por i.

EJEMPLO 4

a. Si w tiene valor absoluto 1, entonces $w = \cos\vartheta + i \operatorname{sen}\vartheta$, donde ϑ es el argumento de w. La multiplicación de cualquier número z diferente de cero por w, simplemente gira z a través de un ángulo ϑ.
b. El argumento de la propia i es $\pi/2$ radianes, así que la multiplicación de z por i gira z a través de un ángulo de $\pi/2$ radianes. Por ejemplo, $3 + i$ gira a $(3 + i)i = -1 + 3i$.

Potencias de un número complejo

La fórmula (4) se usa cuando $z = w = r(\cos\varphi + i \operatorname{sen}\varphi)$. En este caso,

$$z^2 = r^2(\cos 2\varphi + i \operatorname{sen} 2\varphi)$$

y

$$\begin{aligned} z^3 &= z \cdot z^2 \\ &= r(\cos\varphi + i \operatorname{sen}\varphi) \cdot r^2(\cos 2\varphi + i \operatorname{sen} 2\varphi) \\ &= r^3(\cos 3\varphi + i \operatorname{sen} 3\varphi) \end{aligned}$$

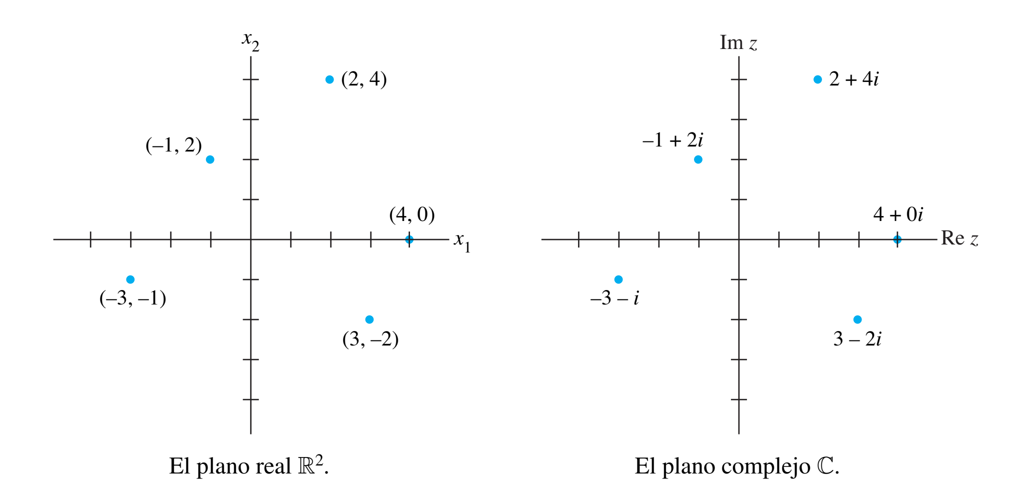
En general, para cualquier entero positivo k,

$$z^k = r^k(\cos k\varphi + i \operatorname{sen} k\varphi)$$

Este hecho se conoce como *Teorema de De Moivre.*

Números complejos y $\mathbb{R}^2$

Aunque los elementos de $\mathbb{R}^2$ y $\mathbb{C}$ tienen una correspondencia uno a uno, y las operaciones de suma son esencialmente las mismas, existe una distinción lógica entre $\mathbb{R}^2$ y $\mathbb{C}$. En $\mathbb{R}^2$ sólo se puede multiplicar un vector por un escalar real, mientras que en $\mathbb{C}$ se puede multiplicar cualesquiera dos números complejos para obtener un tercer número complejo. (El producto punto en $\mathbb{R}^2$ no cuenta, porque produce un escalar, no un elemento de $\mathbb{R}^2$.) Se utiliza la notación escalar para los elementos de $\mathbb{C}$ con el objeto de enfatizar esta distinción.



El plano real $\mathbb{R}^2$. El plano complejo $\mathbb{C}$.

Glosario

A

adjunta (o **adjunta clásica**): Matriz adj A formada a partir de una matriz cuadrada A donde se reemplaza la entrada (i, j) de A por el cofactor (i, j), para todas i y j, y transponiendo después la matriz resultante.

algoritmo de reducción por filas: Método sistemático que utiliza operaciones elementales de fila para reducir una matriz a la forma escalonada o a la forma escalonada reducida.

ampliación de columna-fila: Es la expresión de un producto AB como una suma de productos externos: $\text{col}_1(A)\text{fila}_1(B) + \cdots + \text{col}_n(A)\text{fila}_n(B)$, donde n es el número de columnas de A.

análisis de tendencia: Uso de polinomios ortogonales para ajustar datos, con el producto interior dado mediante la evaluación en un conjunto finito de puntos.

ángulo (entre los vectores distintos de cero **u** y **v** en $\mathbb{R}^2$ o $\mathbb{R}^3$): El ángulo ϑ que se forma entre dos segmentos de recta dirigidos desde el origen hasta los puntos **u** y **v**. Se relaciona con el producto escalar por medio de

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\vartheta$$

aproximación de Fourier (de orden n): Es el punto más cercano a una función dada en $C[0, 2\pi]$, en el subespacio de polinomios trigonométricos de orden n.

aproximación óptima: En un subespacio dado, es el punto más cercano a un vector específico.

aritmética de punto flotante: Aritmética con números representados como decimales $\Delta.d_1 \cdots d_p \,\Delta\, 10^r$, donde r es un entero y la cantidad p de dígitos a la derecha del punto decimal se encuentra usualmente entre 8 y 16.

atractor (de un sistema dinámico en $\mathbb{R}^2$): Es el origen cuando todas las trayectorias tienden a $\mathbf{0}$.

B

base (para un subespacio no trivial H de un espacio vectorial V): Conjunto indexado $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en V tal que: (i) B es un conjunto linealmente independiente, y (ii) el subespacio generado por B coincide con H, esto es, $H = \text{Gen}\,\{\mathbf{v}, \ldots, \mathbf{v}_p\}$.

base de vectores propios: Base completamente constituida por vectores propios de una matriz dada.

base estándar: La base $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ para $\mathbb{R}^n$ consistente en las columnas de la matriz identidad $n \,\Delta\, n$, o la base $\{1, t, \ldots, t^n\}$ para $\mathbb{P}_n$.

base ortogonal: Una base que también es un conjunto ortogonal.

base ortonormal: Base que es un conjunto ortogonal de vectores unitarios.

C

cadena de Markov: Sucesión de vectores de probabilidad $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2, \ldots$, junto con una matriz estocástica P tal que $\boldsymbol{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, 2, \ldots$

cambio de base: *Vea* matriz de cambio de coordenadas.

cambio relativo o **error relativo** (en **b**): La cantidad $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$ cuando **b** cambia a $\mathbf{b} + \Delta\mathbf{b}$.

cociente de Rayleigh: $R(\mathbf{x})$: $(\mathbf{x}^TA\mathbf{x})/\mathbf{x}^T\mathbf{x})$. Estimación de un valor propio de A (usualmente una matriz simétrica).

codominio (de una transformación $T : \mathbb{R}^n \to \mathbb{R}^m$): El conjunto $\mathbb{R}^m$ que contiene el rango de T. En general, si T es función de un espacio vectorial V en un espacio vectorial W, entonces W se denomina codominio de T.

coeficientes de Fourier: Pesos usados para convertir un polinomio trigonométrico en una aproximación de Fourier a una función.

coeficientes de regresión: Son los coeficientes β_0 y β_1 encontrados en la línea de mínimos cuadrados $y = \beta_0 + \beta_1 x$.

cofactor: Es un número $C_{ij} = (-1)^{i+j} \det A_{ij}$, llamado cofactor (i, j) de A, donde A_{ij} es la submatriz que se forma al borrar la i-ésima fila y la j-ésima columna de A.

columna pivote: Columna que contiene una posición pivote.

combinación lineal: Suma de múltiplos escalares de vectores. Los escalares se denominan *pesos*.

complemento de Schur: Cierta matriz formada con bloques de una matriz partida de 2×2 $A = [A_{ij}]$. Si A_{11} es invertible, su complemento de Schur está dado por $A_{22} - A_{21}A_{11}^{-1}A_{12}$.

Si A_{22} es invertible, su complemento de Schur está dado por $A_{11} - A_{12}A_{22}^{-1}A_{21}$.

complemento ortogonal (de W): El conjunto $W^{\perp}$ de todos los vectores ortogonales a W.

componente de y ortogonal a u (para $\mathbf{u} \neq \mathbf{0}$): Es el vector $\mathbf{y} - \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$.

componentes principales (de los datos en una matriz B de observaciones): son los vectores propios unitarios de una matriz de covarianza de muestras S para B, con los vectores propios ordenados de tal forma que los correspondientes valores propios de S disminuyen en magnitud. Si B está en forma de desviación media, entonces los componentes principales son los vectores singulares derechos en una descomposición de valores singulares de B^T.

composición de transformaciones lineales: Función producida al aplicar dos o más transformaciones lineales sucesivas. Si las transformaciones son transformaciones matriciales, por ejemplo la multiplicación izquierda por B seguida por la multiplicación izquierda por A, entonces la composición es $\mathbf{x} \mapsto A(B\mathbf{x})$.

conformables para multiplicación en bloque: Dos matrices partidas A y B tales que el producto en bloque AB está definido: la partición de columnas de A debe coincidir con la partición de filas de B.

conjunto fundamental de soluciones: Es una base para el conjunto de todas las soluciones de una ecuación en diferencias o de una ecuación diferencial lineal homogénea.

conjunto generado por $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$**:** Es el conjunto Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

conjunto generado por $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$**:** Es el conjunto Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

conjunto máximo linealmente independiente (en V): Conjunto B linealmente independiente en V tal que si un vector $\mathbf{v}$ en V pero no en B se agrega a B, entonces el nuevo conjunto es linealmente dependiente.

conjunto mínimo generador (para un subespacio H): Conjunto B que genera a H y tiene la propiedad de que si uno de los elementos de B se retira de B, entonces el nuevo conjunto no genera a H.

conjunto ortogonal: Conjunto de vectores S tal que $\mathbf{u}\cdot\mathbf{v} = 0$ para cada par distinto $\mathbf{u}$, $\mathbf{v}$ en S.

conjunto ortonormal: Es un conjunto ortogonal de vectores unitarios.

conjunto solución: Se compone de todas las soluciones posibles para un sistema lineal. El conjunto solución está vacío cuando el sistema lineal es inconsistente.

contracción: Función $\mathbf{x} \mapsto r\mathbf{x}$ para algún escalar r, con $0 \leq r \leq 1$.

controlable (par de matrices): Par de matrices (A, B) donde A es $n \times n$, B tiene n filas, y rango$[B \quad AB \quad A^2B \quad \cdots \quad A^{n-1}B] = n$ Relacionado con un modelo en el espacio de estados de un sistema de control y la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k$ $(k = 0, 1, \ldots)$.

convergente (sucesión de vectores): Sucesión $[\mathbf{x}_k]$ tal que las entradas en $\mathbf{x}_k$ puedan hacerse tan cercanas como se desee a las entradas en algún vector fijo para toda k lo suficientemente grande.

coordenadas $\mathcal{B}$ de x: *Vea* coordenadas de **x** relativas a la base B.

coordenadas de x relativas a la base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$: Los pesos $c_1, \ldots, c_n$ en la ecuación $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$.

coordenadas homogéneas: En $\mathbb{R}^3$, la representación de (x, y, z) como (X, Y, Z, H) para cualquier $H \neq 0$, donde $x = X/H$, $y = Y/H$ y $z = Z/H$. En $\mathbb{R}^2$, usualmente H se toma como 1, y las coordenadas homogéneas de (x, y) se escriben como $(x, y, 1)$.

corriente de circuito: Cantidad de la corriente eléctrica que fluye a través de un circuito e iguala la suma algebraica de las caídas de voltaje RI alrededor del circuito con la suma algebraica de las fuentes de voltaje del circuito.

covarianza (de las variables x_i y x_j, para $i \neq j$): La entrada s_{ij} en la matriz de covarianza S para una matriz de observaciones, donde x_i y x_j varían con las coordenadas i-ésima y j-ésima, respectivamente, de los vectores de observación.

D

demandas intermedias: Demandas de bienes y servicios que serán consumidos en el proceso de producir otros bienes y servicios para los consumidores. Si **x** es el nivel de producción y C la matriz de consumo, entonces $C\mathbf{x}$ enlista las demandas intermedias.

desarrollo por cofactor: Fórmula para det A que utiliza cofactores asociados con una fila o una columna, tal como para la fila 1: $\det A = a_{11}C_{11} + \cdots + a_{1n}C_{1n}$

desarrollo por cofactores: *Vea* desarrollo por cofactor.

descomposición de vector propio (de **x**): Ecuación, $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, que expresa a **x** como una combinación lineal de los vectores propios de una matriz.

descomposición en valores singulares (de una matriz A $m \times n$): $A = U\Sigma V^T$, donde U es una matriz ortogonal $m \times m$, V una matriz ortogonal $n \,\Delta\, n$, y Σ una matriz $m \,\Delta\, n$ sin entradas negativas en la diagonal principal (establecida en orden de magnitud descendente) y con ceros en todas las demás posiciones. Si el rango $A = r$, entonces Σ tiene exactamente r entradas positivas (los valores singulares distintos de cero en A) sobre la diagonal.

descomposición en valores singulares reducida: Factorización $A = UDV^T$, para una matriz A $m \times n$ con rango r, donde U es $m \times r$ con columnas ortonormales, D es una matriz diagonal $r \times r$ cuyos r valores singulares de A distintos de cero están sobre su diagonal, y V es $n \times r$ con columnas ortonormales.

descomposición espectral (de A): Representación

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T$$

donde $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ es una base ortonormal de vectores propios de A, y $\lambda_1, \ldots, \lambda_n$ son los correspondientes valores propios de A.

descomposición ortogonal: Es la representación de un vector $\mathbf{y}$ como la suma de dos vectores, uno en un subespacio específico W y el otro en $W^\perp$. En general, una descomposición $\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$, donde $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una base ortogonal para un subespacio que contiene a $\mathbf{y}$.

descomposición polar (de A): Factorización $A = PQ$, donde P es una matriz semidefinida positiva $n \times n$ con el mismo rango que A, y Q es una matriz ortogonal $n \times n$.

descripción explícita (de un subespacio W de $\mathbb{R}^n$): Representación paramétrica de W como el conjunto de todas las combinaciones lineales de un conjunto de vectores específico.

descripción implícita (de un subespacio W de $\mathbb{R}^n$): Conjunto de una o más ecuaciones homogéneas que caracterizan los puntos de W.

desigualdad de Cauchy-Schwarz: $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|u\|\cdot\|v\|$ para toda $\mathbf{u}$ y $\mathbf{v}$.

desigualdad triangular: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ para toda $\mathbf{u}$ y $\mathbf{v}$.

determinante (de una matriz cuadrada A): Es el número $\det A$ definido en forma inductiva mediante un desarrollo por cofactores a lo largo de la primera fila de A. También, $(-1)^r$ veces el producto de las entradas diagonales en cualquier forma escalonada U obtenida a partir de A mediante reemplazos de fila y r intercambios de fila (pero sin operaciones de escalamiento).

diagonal en bloque (matriz): Es una matriz partida $A = [A_{ij}]$, tal que cada bloque A_{ij} es una matriz cero para $i \neq j$.

diagonal principal (de una matriz): Las entradas con iguales índices de fila y columna.

diagonalizable (matriz): Matriz que puede escribirse en forma factorizada como PDP^{-1}, donde D es una matriz diagonal y P es una matriz invertible.

dilatación: Función $\mathbf{x} \mapsto r\mathbf{x}$ para algún escalar r, con $1 < r$.

dimensión (de un espacio vectorial V): Es el número de vectores presentes en una base de V, se escribe como $\dim V$. La dimensión del espacio cero es 0.

dimensional finito (espacio vectorial): Un espacio vectorial generado por un conjunto finito de vectores.

dimensional infinito (espacio vectorial): Espacio vectorial V distinto de cero que tiene una base no finita.

distancia a un subespacio: La distancia medida desde un punto (vector) dado $\mathbf{v}$ hasta el punto más cercano en el subespacio.

distancia entre u y v: Longitud del vector $\mathbf{u} - \mathbf{v}$, denotada mediante $\text{dist}\,(\mathbf{u}, \mathbf{v})$.

distinto de cero (matriz o vector): Matriz (posiblemente con sólo una fila o columna) que contiene al menos una entrada distinta de cero.

dominio (de una transformación T): Es el conjunto de todos los vectores $\mathbf{x}$ para los cuales $T(\mathbf{x})$ está definida.

E

ecuación auxiliar: Es una ecuación polinomial en una variable r, se crea a partir de los coeficientes de una ecuación en diferencias homogénea.

ecuación característica (de A): $\det(A - \lambda I) = 0$.

ecuación en diferencias (o **relación de recurrencia lineal**): Ecuación de la forma $\mathbf{x}_{\mathbf{k+1}} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$ cuya solución es una sucesión de vectores $\mathbf{x}_0, \mathbf{x}_1, \ldots$.

ecuación homogénea: Ecuación de la forma $A\mathbf{x} = \mathbf{0}$, posiblemente escrita como una ecuación de vector o como un sistema de ecuaciones lineales.

ecuación lineal (en las variables $x_1, \ldots, x_n$): Expresión que puede escribirse en la forma $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, donde b y los coeficientes $a_1, \ldots, a_n$ son números reales o complejos.

ecuación matricial: Expresión que involucra al menos una matriz; por ejemplo, $A\mathbf{x} = \mathbf{b}$.

ecuación no homogénea: Expresión de la forma $A\mathbf{x} = \mathbf{b}$ con $\mathbf{b} \neq \mathbf{0}$, escrita posiblemente como una ecuación vectorial o un sistema de ecuaciones lineales.

ecuación paramétrica de un plano: Expresión del tipo $\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$ (s, t en $\mathbb{R}$), siendo $\mathbf{u}$ y $\mathbf{v}$ linealmente independientes.

ecuación paramétrica de una recta: Expresión del tipo $\mathbf{x} = \mathbf{p} + t\mathbf{v}$ (t en $\mathbb{R}$).

ecuación vectorial: Expresión que involucra una combinación lineal de vectores con pesos no determinados.

ecuaciones normales: Sistema de ecuaciones representado mediante $A^TA\mathbf{x} = A^T\mathbf{b}$, cuya solución produce todas las soluciones de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. En estadística, una notación común es $X^TX\boldsymbol{\beta} = X^T\mathbf{y}$.

ejes principales (de una forma cuadrática $\mathbf{x}^TA\mathbf{x}$): son las columnas ortonormales de una matriz ortogonal P tal que $P^{-1}AP$ es diagonal. (Estas columnas son valores unitarios propios de A.) Por lo general, las columnas de P se ordenan de tal forma que los correspondientes valores propios de A se establecen en orden de magnitud descendente.

eliminación gaussiana: *Vea* algoritmo de reducción por filas.

entrada principal: Es la entrada distinta de cero que se encuentra más a la izquierda de una fila de una matriz.

entradas diagonales (en una matriz): Entradas que tienen índices iguales de fila y columna.

equivalentes de fila (matrices): Dos matrices para las cuales existe una sucesión (finita) de operaciones de fila que transforman una matriz en otra.

error cuadrado medio: Error de una aproximación en un espacio de producto interior, donde el producto interior se define mediante una integral definida.

error de redondeo: En la aritmética de punto flotante, error ocasionado cuando el resultado de un cálculo se redondea (o trunca) al número almacenado de dígitos de punto flotante. También, error que resulta cuando la representación decimal

de un número, tal como 1/3, se aproxima mediante un número de punto flotante con cierta cantidad finita de dígitos.

error por mínimos cuadrados: La distancia $\|\mathbf{b} - A\hat{\mathbf{x}}\|$ de $\mathbf{b}$ a $A\hat{\mathbf{x}}$, cuando $\hat{\mathbf{x}}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

escalamiento (de un vector): Multiplicar un vector (o bien una fila o una columna de una matriz) por un escalar distinto de cero.

escalar: Número (real) usado para multiplicar un vector o una matriz.

espacio con producto interior: Espacio vectorial sobre el que se define un producto interior.

espacio de columna (de una matriz A de $m \times n$): Es el conjunto Col A de todas las combinaciones lineales de las columnas de A. Si $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$, entonces Col $A = \text{Gen}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$. De manera equivalente

$$\text{Col } A = \{\mathbf{y} : \mathbf{y} = A\mathbf{x} \text{ para alguna } \mathbf{x} \text{ en } \mathbb{R}^n\}$$

espacio fila (de una matriz A): Conjunto Fila A de todas las combinaciones lineales de los vectores formados con las filas de A; también se denota mediante Col A^T.

espacio nulo (de una matriz A m Δ n): Es el conjunto Nul A de todas las soluciones a la ecuación homogénea $A\mathbf{x} = \mathbf{0}$. Nul $A = \{\mathbf{x} : \mathbf{x} \text{ está en } \mathbb{R}^n \text{ y } A\mathbf{x} = 0\}$.

espacio propio (de A correspondiente a λ): El conjunto de *todas* las soluciones de $A\mathbf{x} = \lambda\mathbf{x}$, donde λ es un valor propio de A. Consta del vector cero y de todos los vectores propios correspondientes a λ.

espacio vectorial: Es un conjunto de objetos, llamados vectores, sobre los que se definen dos operaciones, denominadas suma y multiplicación por escalares. Deben satisfacerse diez axiomas. Vea la primera definición en la sección 4.1.

espacios vectoriales isomorfos: Dos espacios vectoriales V y W para los cuales existe una transformación lineal T uno a uno que mapea a V sobre W.

F

factorización (de A): Ecuación que expresa A como un producto de dos o más matrices.

factorización de Cholesky: Es una factorización $A = R^TR$, donde R es una matriz triangular superior invertible cuyas entradas diagonales son todas positivas.

factorización de Schur (de A, para escalares reales): Una factorización $A = URU^T$ de una matriz A $n \times n$ que tiene n valores propios reales, donde U es una matriz ortogonal $n \times n$ y R una matriz triangular superior.

factorización LU permutada: Representación de una matriz A en la forma $A = LU$, donde L es una matriz cuadrada tal que una permutación de sus filas formará una matriz triangular inferior unitaria, y U es una forma escalonada de A.

factorización LU: Representación de una matriz A en la forma $A = LU$, donde L es una matriz triangular inferior cuadrada con números uno en la diagonal (una matriz triangular inferior unitaria), y U es una forma escalonada de A.

factorización QR: Factorización de una matriz A m Δ n cuyas columnas forman una base ortonormal para Col A, y R es una matriz invertible triangular superior n Δ n con entradas positivas sobre su diagonal.

fase progresiva (de la reducción por filas): Primera parte del algoritmo que reduce una matriz a la forma escalonada.

fase regresiva (de la reducción por filas): Última parte del algoritmo que reduce una matriz en forma escalonada a una forma escalonada reducida.

filtro lineal: Ecuación lineal en diferencias usada para transformar señales de tiempo discreto.

flop: Operación aritmética (+, −, *, /) sobre dos números reales de punto flotante.

forma cuadrática definida negativa: Expresión cuadrática Q tal que $Q(\mathbf{x}) < 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

forma cuadrática definida positiva: Expresión cuadrática Q tal que $Q(\mathbf{x}) > 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

forma cuadrática indefinida: Expresión cuadrática Q tal que $Q(\mathbf{x})$ asume valores tanto positivos como negativos.

forma cuadrática semidefinida negativa: Forma cuadrática Q tal que $Q(\mathbf{x}) \leq 0$ para toda $\mathbf{x}$.

forma cuadrática semidefinida positiva: Es un forma cuadrática Q tal que $Q(\mathbf{x}) \geq 0$ para toda $\mathbf{x}$.

forma cuadrática: Función Q definida para $\mathbf{x}$ en $\mathbb{R}^n$ mediante $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, donde A es una matriz simétrica $n \times n$ (llamada **matriz de la forma cuadrática**).

forma de desviación media (de un vector): Vector cuyas entradas suman cero.

forma de desviación media (de una matriz de observaciones): Matriz cuyos vectores fila están en forma de desviación media. Para cada fila, las entradas suman cero.

forma escalonada (o **forma escalonada por filas**, de una matriz): Matriz escalonada que es equivalente por filas a la matriz dada.

forma escalonada reducida (o **forma escalonada reducida por filas**): Matriz escalonada reducida que es equivalente por filas a una matriz dada.

función (o mapeo): *Vea* transformación.

función de coordenadas (determinada mediante una base ordenada $\mathcal{B}$ en un espacio vectorial V): Función que asocia a cada $\mathbf{x}$ en V su vector coordenado $[\mathbf{x}]_{\mathcal{B}}$.

funciones propias (de una ecuación diferencial $\mathbf{x}'(t) = A\mathbf{x}(t)$): Funciones del tipo $\mathbf{x}(t) = \mathbf{v}e^{\lambda t}$, donde $\mathbf{v}$ es un vector propio de A yλ es el valor propio correspondiente.

G

gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$: Conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$. También, el *subespacio generado* por $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

I

Im x: Vector en $\mathbb{R}^n$ formado con las partes imaginarias de las entradas de un vector **x** en $\mathbb{C}^n$.

imagen (de un vector **x** bajo una transformación T): Es el vector $T(\mathbf{x})$ asignado a **x** por medio de T.

inversa (de una matriz A $n \times n$): Matriz A^{-1} $n \times n$ tal que $AA^{-1} = A^{-1}A = I_n$.

inversa de Moore-Penrose: *Vea* seudoinverso.

inversa derecha (de A): Cualquier matriz rectangular C tal que $AC = I$.

inversa izquierda (de A): Cualquier matriz rectangular C tal que $CA = I$.

isomorfismo: Transformación lineal uno a uno de un espacio vectorial sobre otro.

L

ley asociativa de la multiplicación: $A(BC) = (AB)C$, para toda A, B y C.

leyes de Kirchhoff: (1) (**ley del voltaje**) La suma algebraica de las caídas de voltaje RI en una dirección alrededor de un circuito es igual a la suma algebraica de las fuentes de voltaje que van en la misma dirección alrededor del circuito. (2) (**ley de la corriente**) La corriente en una rama es la suma algebraica de las corrientes del circuito que fluyen a través de dicha rama.

leyes distributivas: (izquierda) $A(B + C) = AB + AC$, y (derecha) $(B + C)A = BA + CA$, para toda A, B y C.

línea a través de p paralela a v: El conjunto $\{\mathbf{p} + t\mathbf{v}\text{: t en } \mathbb{R}\}$.

línea de mínimos cuadrados: La línea línea $y = \hat{\beta}_0 + \hat{\beta}_1 x$ que minimiza el error de mínimos cuadrados en la ecuación $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$.

linealmente dependientes (vectores): Conjunto indexado $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ con la propiedad de que existen los pesos $c_1, \ldots, c_p$, no todos iguales a cero, tales que $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$. Esto es, la ecuación vectorial $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$ tiene una solución *no trivial*.

linealmente independientes (vectores): Conjunto indexado $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ con la propiedad de que la ecuación vectorial $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$ tiene *únicamente* la solución trivial, $c_1 = \cdots = c_p = 0$.

longitud (o **norma**, de **v**): Es el escalar $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$.

M

magnitud (de un vector): *Vea* norma.

matrices conmutativas: Dos matrices A y B tales que $AB = BA$.

matriz aumentada: Matriz que se estructura mediante una matriz de coeficientes para un sistema lineal y una o más columnas a la derecha. Cada columna extra contiene las constantes del lado derecho de un sistema con la matriz de coeficientes dada.

matriz bidiagonal: Matriz cuyas entradas distintas de cero pertenecen a la diagonal principal y a una diagonal adyacente a la principal.

matriz casi singular: Matriz mal condicionada.

matriz compañera: Forma especial de matriz cuyo polinomio característico es $(-1)^n p(\lambda)$ cuando $p(\lambda)$ es un polinomio específico cuyo término principal es λ^n.

matriz de bandas: Es una matriz cuyas entradas distintas de cero están dentro de una banda situada a lo largo de la diagonal principal.

matriz de cambio de coordenadas (de una base $\mathcal{B}$ a una base $\mathcal{C}$): Matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ que transforma vectores de coordenadas $\mathcal{B}$ en vectores de coordenadas $\mathcal{C}$: $[\mathbf{x}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}}$. Si $\mathcal{C}$ es la base estándar para $\mathbb{R}^n$, entonces $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ se escribe ocasionalmente como $P_{\mathcal{B}}$.

matriz de coeficientes: Matriz cuyas entradas son los coeficientes de un sistema de ecuaciones lineales.

matriz de consumo: En el modelo de entrada y salida de Leontief, matriz cuyas columnas son los vectores de consumo unitarios para los diferentes sectores de una economía.

matriz de covarianza (o **matriz de covarianza muestral**)**:** Es la matriz S de $p \;\Delta\; p$ definida mediante $S = (N - 1)^{-1}BB^T$, donde B es una matriz de observaciones $p \times N$ en la forma de desviación media.

matriz de diseño: Es la matriz X en el modelo lineal $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde las columnas de X son determinadas en cierta forma por los valores observados de algunas variables independientes.

matriz de entrada y salida: *Vea* matriz de consumo.

matriz de flexibilidad: Matriz cuya j-ésima columna proporciona las deflexiones de una viga elástica en puntos específicos cuando una fuerza unitaria se aplica en el j-ésimo punto de la viga.

matriz de Gram (de A): La matriz A^TA.

matriz de migración: Matriz que proporciona el movimiento porcentual entre diferentes ubicaciones, de un periodo al próximo.

matriz de observaciones: Matriz $p \;\Delta\; N$ cuyas columnas son vectores de observación, cada columna enlista p mediciones hechas sobre un individuo u objeto en una población específica o un conjunto dado.

matriz de proyección (o **matriz de proyección ortogonal**)**:** Matriz simétrica B tal que $B^2 = B$. Un ejemplo simple es $B = \mathbf{v}\mathbf{v}^T$, donde **v** es un vector unitario.

matriz de rigidez: Inversa de una matriz de flexibilidad. La j-ésima columna de una matriz de rigidez proporciona las cargas que deben aplicarse en puntos específicos de una viga elástica para producir una deflexión unitaria en el j-ésimo punto de la viga.

matriz de transferencia: Matriz A asociada con un circuito eléctrico que tiene terminales de entrada y salida, de modo que el vector de salida es A veces el vector de entrada.

matriz de Vandermonde: Matriz V $n \times n$ o su transpuesta, cuando V tiene la forma

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

matriz definida negativa: Matriz simétrica A tal que $\mathbf{x}^T A\mathbf{x} < 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

matriz definida positiva: Matriz simétrica A tal que $\mathbf{x}^T A\mathbf{x} > 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

matriz diagonal: Matriz cuadrada cuyas entradas que *no* se encuentran en la diagonal principal son iguales a cero.

matriz elemental: Matriz invertible que resulta de efectuar una operación elemental de fila sobre una matriz identidad.

matriz en bloque: *Vea* matriz partida.

matriz escalonada (o **matriz escalonada por filas):** Matriz rectangular que tiene tres propiedades: (1) Todas las filas distintas de cero se encuentran arriba de cualquier fila de ceros. (2) Cada entrada principal de una fila se encuentra en una columna a la derecha de la entrada principal de la fila que está arriba de ésta. (3) Todas las entradas ubicadas en una columna por debajo de una entrada principal son iguales a cero.

matriz escalonada reducida: Matriz rectangular en forma escalonada que tiene estas propiedades adicionales: la entrada principal en cada fila distinta de cero es 1, y cada 1 delantero es la única entrada distinta de cero localizada en su columna.

matriz estándar (para una transformación lineal T): Es una matriz A tal que $T(\mathbf{x}) = A\mathbf{x}$ para toda $\mathbf{x}$ en el dominio de T.

matriz estocástica regular: Matriz estocástica P tal que alguna potencia de la matriz P^k contiene sólo entradas estrictamente positivas.

matriz estocástica: Es una matriz cuadrada cuyas columnas son vectores de probabilidad.

matriz identidad (denotada mediante I o I_n): Matriz cuadrada con números uno sobre la diagonal y ceros en todas las demás posiciones.

matriz indefinida: Es una matriz simétrica A tal que $\mathbf{x}^T A\mathbf{x}$ asume tanto valores positivos como negativos.

matriz invertible: Matriz cuadrada que posee una inversa.

matriz $m \times n$: Una matriz con m filas y n columnas.

matriz mal condicionada: Matriz cuadrada con un número de condición muy grande (o posiblemente infinito); matriz que es singular o puede convertirse en singular si algunas de sus entradas se cambian un poco.

matriz ortogonal: Matriz cuadrada invertible U tal que $U^{-1} = U^T$.

matriz para T relativa a las bases $\mathcal{B}$ y $\mathcal{C}$: Matriz M para una transformación lineal $T : V \to W$ con la propiedad de que $[T(\mathbf{x})]_{\mathcal{C}} = M[\mathbf{x}]_{\mathcal{B}}$ para toda $\mathbf{x}$ en V, donde $\mathcal{B}$ es una base para V y $\mathcal{C}$ es una base para W. Cuando $W = V$ y $\mathcal{C} = \mathcal{B}$, la matriz M se denomina matriz $\mathcal{B}$ para T y se denota con $[T]_{\mathcal{B}}$.

matriz partida (o **matriz de bloques**): Matriz cuyas entradas son, a su vez, matrices de tamaño adecuado.

matriz semidefinida negativa: Matriz simétrica A tal que $\mathbf{x}^T A\mathbf{x} \leq 0$ para toda $\mathbf{x}$.

matriz semidefinida positiva: Matriz simétrica A tal que $\mathbf{x}^T A\mathbf{x} \geq 0$ para toda $\mathbf{x}$.

matriz simétrica: Es una matriz A tal que $A^T = A$.

matriz triangular inferior permutada: Matriz tal que una permutación de sus filas formará una matriz triangular inferior.

matriz triangular inferior unitaria: Es una matriz triangular inferior cuadrada con números uno sobre la diagonal principal.

matriz triangular inferior: Matriz con ceros arriba de la diagonal principal.

matriz triangular superior: Matriz U (no necesariamente cuadrada) con ceros debajo de las entradas diagonales u_{11}, u_{22}, ...

matriz triangular: Matriz A que tiene ceros ya sea arriba o abajo de las entradas diagonales.

matriz: Arreglo rectangular de números.

matriz-$\mathcal{B}$ (para T): Matriz $[T]_B$ para una transformación lineal T: $V \,\Delta\, V$ relativa a la base B para V, con la propiedad de que $[T(\mathbf{x})]_B = [T]_B[\mathbf{x}]_B$ para toda $\mathbf{x}$ en V.

media de la muestra: Promedio $\mathbf{M}$ de un conjunto de vectores, $\mathbf{X}_1, \ldots, \mathbf{X}_N$, dados por $\mathbf{M} = (1/N)(\mathbf{X}_1 + \cdots + \mathbf{X}_N)$.

método de la potencia inversa: Algoritmo utilizado para estimar un valor propio λ de una matriz cuadrada, cuando existe una buena estimación inicial de λ.

método de potencia: Algoritmo utilizado para estimar un valor propio estrictamente dominante de una matriz cuadrada.

mínimos cuadrados ponderados: Problemas de mínimos cuadrados con un producto interior ponderado como

$$\langle \mathbf{x}, \mathbf{y} \rangle = w_1^2 x_1 y_1 + \cdots + w_n^2 x_n y_n.$$

misma dirección (como un vector $\mathbf{v}$): Vector que es un múltiplo positivo de $\mathbf{v}$.

modelo de entrada y salida de Leontief (o **ecuación de producción de Leontief**): Es la ecuación $\mathbf{x} = C\mathbf{x} + \mathbf{d}$, donde $\mathbf{x}$ es la producción, $\mathbf{d}$ la demanda final, y C la demanda intermedia (o matriz de entrada y salida). La j-ésima columna de C enlista las entradas que consume el sector j por unidad de salida.

modelo de entrada y salida: *Vea* modelo de entrada y salida de Leontief.

modelo de intercambio de Leontief (o **modelo cerrado**): Modelo de una economía donde las entradas y las salidas son fijas, y donde un conjunto de precios para las salidas de los sectores es visto de tal manera que el ingreso de cada sector es igual a los gastos que realiza. Esta condición de "equilibrio" se expresa como un sistema de ecuaciones lineales, siendo los precios las incógnitas.

modelo de intercambio: *Vea* modelo de intercambio de Leontief.

modelo lineal (en estadística): Cualquier ecuación de la forma $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\Delta}$, donde X e $\mathbf{y}$ son conocidas y $\boldsymbol{\beta}$ se elige para minimizar la longitud del **vector residual** $\boldsymbol{\epsilon}$.

modelo matricial por etapas: Ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde $\mathbf{x}_k$ enlista la cantidad de mujeres incluidas en una población en el tiempo k, con las mujeres clasificadas en diferentes etapas de desarrollo (tal como juvenil, subadulta y adulta).

multiplicación de matriz en bloque: Multiplicación fila-columna de matrices partidas como si las entradas de bloque fueran escalares.

multiplicación derecha (por A): Multiplicación por A de una matriz sobre la derecha.

multiplicación izquierda (por A): Multiplicación por A de un vector o una matriz sobre el lado izquierdo.

multiplicidad algebraica: Diversidad de un valor propio como una raíz de la ecuación característica.

múltiplo escalar de u por ***c*****:** Es el vector $c\mathbf{u}$ obtenido al multiplicar cada entrada en $\mathbf{u}$ por c.

N

no singular (matriz): Es una matriz invertible.

norma (o **longitud**, de $\mathbf{v}$): Es el escalar $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{\langle\mathbf{v}, \mathbf{v}\rangle}$.

normalización (de un vector $\mathbf{v}$ distinto de cero): Proceso de creación de un vector unitario $\mathbf{u}$ que es un múltiplo positivo de $\mathbf{v}$.

núcleo (de una transformación lineal $T: V \to W$): Conjunto de las $\mathbf{x}$ en V tales que $T(\mathbf{x}) = \mathbf{0}$.

número de condición (de A): Es el cociente σ_1/σ_n, donde σ_1 es el valor singular máximo de A y σ_n es el valor singular mínimo. El número de condición es $+\infty$ cuando σ_n es cero.

O

operaciones elementales de fila: (1) (Reemplazo) Sustituir una fila con la suma de esa fila y de un múltiplo perteneciente a otra fila. (2) Intercambio de dos filas. (3) (Escalamiento) Multiplicar todas las entradas de una fila por una constante distinta de cero.

optimización restringida: El problema de maximizar una cantidad como $\mathbf{x}^TA\mathbf{x}$ o $\|A\mathbf{x}\|$ cuando $\mathbf{x}$ está sujeta a una o más restricciones, tales como $\mathbf{x}^T\mathbf{x} = 1$ o $\mathbf{x}^T\mathbf{v} = 0$.

origen: Es el vector cero.

ortogonal a ***W*****:** Ortogonal a cualquier vector en W.

ortogonalmente diagonalizable (matriz): Matriz A que admite una factorización, $A = PDP^{-1}$, con P como una matriz ortogonal ($P^{-1} = P^T$) y diagonal D.

P

parte triangular inferior (de A): Matriz triangular inferior cuyas entradas en la matriz diagonal y debajo de ella coinciden con las entradas de A.

pesos: Escalares usados en una combinación lineal.

pivote: Número distinto de cero que se utiliza en una posición pivote para crear ceros mediante operaciones de fila o que se transforma en un 1 principal, empleado a su vez para crear ceros.

plano a través de u, v y el origen: Conjunto cuya ecuación paramétrica es $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ (s, t en $\mathbb{R}$), siendo $\mathbf{u}$ y $\mathbf{v}$ linealmente independientes.

polinomio característico (de A): $\det(A - \lambda I)$ o, en algunos textos, $\det(\lambda I - A)$.

polinomio de interpolación: Expresión cuya gráfica pasa por todos los puntos de un conjunto de datos en $\mathbb{R}^2$.

polinomio trigonométrico: Combinación lineal compuesta por la función constante 1 y las funciones seno y coseno tales como cos nt y sen nt.

posición estándar: La posición de la gráfica de una ecuación $\mathbf{x}^TA\mathbf{x} = c$ cuando A es una matriz diagonal.

posición pivote: En una matriz A, posición que corresponde a una entrada principal en una forma escalonada de A.

precios de equilibrio: Conjunto de precios para la producción total de los diferentes sectores de una economía, tal que el ingreso de cada sector balancee de manera exacta sus gastos.

pregunta de existencia: Pregunta, "¿Existe una solución al sistema?" Esto es, "¿El sistema es consistente?" También, "¿Existe una solución de $A\mathbf{x} = \mathbf{b}$ para *todas* las $\mathbf{b}$ posibles?"

pregunta de unicidad: Pregunta, "Si existe una solución a un sistema, ¿es única?; esto es, ¿sólo existe una solución?"

problema general de mínimos cuadrados: Dados una matriz A de $m \times n$ y un vector $\mathbf{b}$ en $\mathbb{R}^m$, encuentre $\hat{\mathbf{x}}$ en $\mathbb{R}^n$ tal que $\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

proceso de Gram-Schmidt: Algoritmo empleado para producir una base ortogonal u ortonormal para un subespacio generado por un conjunto dado de vectores.

producto ***A*****x:** Combinación lineal de las columnas de A usando las correspondientes entradas en $\mathbf{x}$ como pesos.

producto exterior: Producto matricial $\mathbf{u}\mathbf{v}^T$ donde $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$ vistos como matrices $n \times 1$. (El símbolo de transpuesta está en el "exterior" de los símbolos $\mathbf{u}$ y $\mathbf{v}$.)

producto interior: Es el escalar $\mathbf{u}^T\mathbf{v}$, que usualmente se escribe como $\mathbf{u}\cdot\mathbf{v}$, donde $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$ vistos como matrices de $n \times 1$. También es llamado **producto punto** de $\mathbf{u}$ y $\mathbf{v}$. En general, es una función sobre un espacio vectorial que asigna a cada par de vectores $\mathbf{u}$ y $\mathbf{v}$ un número $\langle\mathbf{u}, \mathbf{v}\rangle$, sujeto a ciertos axiomas. Vea la sección 6.7.

producto punto: *Vea* producto interior.

proyección ortogonal de y sobre u (o sobre la línea que pasa por $\mathbf{u}$ y el origen, para $\mathbf{u} \neq \mathbf{0}$): El vector $\hat{\mathbf{y}}$ definido mediante $\hat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$.

proyección ortogonal de y sobre ***W*****:** El vector único $\hat{\mathbf{y}}$ en W tal que $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a W. Notación: $\hat{\mathbf{y}} = \text{proy}_W\, \mathbf{y}$.

punto espiral (de un sistema dinámico en $\mathbb{R}^2$): Es el origen cuando las trayectorias son espirales alrededor de **0**.

punto silla (de un sistema dinámico en $\mathbb{R}^2$): Es el origen cuando algunas trayectorias son atraídas hacia **0** y otras trayectorias son repelidas desde **0**.

R

rango (de una matriz A): Dimensión del espacio columna de A, denotado mediante rango A.

rango (de una transformación lineal T): Conjunto de todos los vectores de la forma $T(\mathbf{x})$ para alguna $\mathbf{x}$ en el dominio de T.

rango pleno (matriz): Matriz $m \times n$ cuyo rango es el menor de m y n.

Re x: Vector en $\mathbb{R}^n$ formado con las partes reales de las entradas de un vector $\mathbf{x}$ en $\mathbb{C}^n$.

red en escalera: Red eléctrica ensamblada mediante la conexión en serie de dos o más circuitos eléctricos.

reemplazo de fila: Operación elemental de fila que sustituye una fila de una matriz por la suma de dicha fila y de un múltiplo de otra fila.

reflexión de Householder: Transformación $< \mathbf{x} \mapsto Q\mathbf{x}$, donde $Q = I - 2\mathbf{u}\mathbf{u}^T$ y $\mathbf{u}$ es un vector unitario ($\mathbf{u}^T\mathbf{u} = 1$).

regla de Cramer: Fórmula para cada entrada en la solución $\mathbf{x}$ de la ecuación $A\mathbf{x} = \mathbf{b}$ cuando A es una matriz invertible.

regla del paralelogramo para la suma: Interpretación geométrica de la suma de dos vectores $\mathbf{u}$ y $\mathbf{v}$ como la diagonal del paralelogramo determinado mediante $\mathbf{u}$, $\mathbf{v}$ y **0**.

regla fila-columna: Pauta para calcular un producto AB en el que la entrada (i, j) de AB es la suma de los productos de entradas correspondientes desde la fila i de A y la columna j de B.

regla fila-vector para calcular $A\mathbf{x}$: Pauta para calcular un producto $A\mathbf{x}$ en el que la i-ésima entrada de $A\mathbf{x}$ es la suma de los productos de las entradas correspondientes de la fila i de A y del vector $\mathbf{x}$.

regresión múltiple: Modelo lineal que involucra algunas variables independientes y una variable dependiente.

relación de dependencia lineal: Ecuación vectorial homogénea donde se especifican todos los pesos y al menos un peso es diferente de cero.

relación de recurrencia: *Vea* ecuación en diferencias.

Relacionado con un modelo en el espacio de estados de un sistema de control y la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k$ ($k = 0, 1, \ldots$).

repulsor (de un sistema dinámico en $\mathbb{R}^2$): Es el origen cuando todas las trayectorias, excepto la sucesión o función de la constante cero, tienden a alejarse de **0**.

resta vectorial: Cálculo de $\mathbf{u} + (-1)\mathbf{v}$ y escritura del resultado como $\mathbf{u} - \mathbf{v}$.

rotación de Givens: Transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^n$ utilizada en programas de computadora para crear entradas cero en un vector (casi siempre una columna de una matriz).

S

señal (o **señal de tiempo discreto**)**:** Sucesión doblemente infinita de números $\{y_k\}$; una función definida en los enteros; pertenece al espacio vectorial $\mathbb{S}$.

serie de Fourier: Secuencia infinita que converge hacia una función en el espacio del producto interior $C[0, 2\pi]$, con el producto interior dado por una integral definida.

seudoinverso (de A): Es la matriz $VD^{-1}U^T$, cuando UDV^T es una descomposición del valor singular reducido de A.

similares (matrices): Matrices A y B tales que $P^{-1}AP = B$, o de manera equivalente, $A = PBP^{-1}$, para alguna matriz invertible P.

singular (matriz): Matriz cuadrada que no tiene inversa.

sistema de ecuaciones lineales (o **sistema lineal**): Colección de una o más ecuaciones lineales que involucran al mismo conjunto de variables, por ejemplo, $x_1, \ldots, x_n$.

sistema desacoplado: Ecuación en diferencias $\mathbf{y}_{k+1} = A\mathbf{y}_k$, o bien una ecuación diferencial $\mathbf{y}'(t) = A\mathbf{y}(t)$, en la cual A es una matriz diagonal. La evolución discreta de cada entrada en $\mathbf{y}_k$ (como una función de k), o la evolución continua de cada entrada en la función con valores vectoriales $\mathbf{y}(t)$, no se ve afectada por lo que sucede con las otras entradas cuando $k \to \infty$ o $t \to \infty$.

sistema dinámico lineal discreto: Ecuación en diferencias de la forma $\mathbf{x}_{k+1} = A\mathbf{x}_k$ que describe los cambios ocurridos en un sistema (por lo general, un sistema físico) conforme pasa el tiempo. El sistema físico se mide en tiempos discretos, cuando $k = 0, 1, 2, \ldots$, y el **estado** del sistema en un tiempo k es un vector $\mathbf{x}_k$ cuyas entradas dan a conocer ciertos aspectos de interés acerca del sistema.

sistema dinámico: *Vea* sistema dinámico lineal discreto.

sistema lineal consistente: Sistema lineal que tiene al menos una solución.

sistema lineal inconsistente: Sistema lineal sin solución.

sistema lineal: Colección de una o más ecuaciones lineales que involucran las mismas variables, por ejemplo, $x_1, \ldots, x_n$.

sistema sobredeterminado: Sistema de ecuaciones con más ecuaciones que incógnitas.

sistema subdeterminado: Sistema de ecuaciones con menos ecuaciones que incógnitas.

sistemas (lineales) equivalentes: Sistemas lineales con el mismo conjunto solución.

sobre un conjunto (función): Función $T : \mathbb{R}^n \to \mathbb{R}^m$ tal que cada $\mathbf{b}$ en R^m es la imagen de *al menos* una $\mathbf{x}$ en $\mathbb{R}^n$.

solución (de un sistema lineal que incluye las variables $x_1, \ldots, x_n$): Lista de números $(s_1, s_2, \ldots, s_n)$ que hace de cada ecuación presente en el sistema un enunciado verdadero cuando los valores $s_1, \ldots, s_n$ sustituyen, respectivamente a $x_1, \ldots, x_n$.

solución general (de un sistema lineal): Descripción paramétrica de un conjunto solución que expresa las variables básicas en

términos de las variables libres (los parámetros), si existe alguna. Después de la sección 1.5, la descripción paramétrica se escribe en forma vectorial.

solución no trivial: Solución distinta de cero para una ecuación homogénea o un sistema de ecuaciones homogéneas.

solución por mínimos cuadrados (de $A\mathbf{x} = \mathbf{b}$): Vector $\hat{\mathbf{x}}$ tal que $\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

solución trivial: Es la solución $\mathbf{x} = \mathbf{0}$ de una ecuación homogénea $A\mathbf{x} = \mathbf{0}$.

subespacio cero: Es el subespacio $\{\mathbf{0}\}$ que consta sólo del vector cero.

subespacio invariante (para A): Subespacio H tal que $A\mathbf{x}$ está en H siempre que $\mathbf{x}$ esté en H.

subespacio propio: Cualquier subespacio de un espacio vectorial V diferente al mismo V.

subespacio: Subconjunto H de algún espacio vectorial V tal que H tiene estas propiedades: (1) El vector cero de V está en H; (2) H es cerrado bajo la suma de vectores; y (3) H es cerrado bajo la multiplicación por escalares.

subespacios fundamentales (determinados mediante A): El espacio nulo y el espacio de columna de A, y el espacio nulo y el espacio de columna de A^T, donde generalmente a Col A^T se le llama espacio de fila de A.

submatriz (de A): Cualquier matriz obtenida al borrar algunas filas y/o columnas de A; también es la propia A.

suma de columna: Suma de las entradas que haya en una columna de una matriz.

suma de filas: Suma de las entradas presentes en una fila de una matriz.

suma vectorial: Adición de vectores al sumar entradas correspondientes.

sustitución regresiva (con notación matricial): Fase regresiva de la reducción por filas de una matriz aumentada que transforma una matriz escalonada en una matriz escalonada reducida; se usa para encontrar la solución o soluciones de un sistema de ecuaciones lineales.

T

tamaño (de una matriz): Dos números, escritos en la forma $m \times n$, que especifican la cantidad de filas (m) y columnas (n) presentes en la matriz.

término de producto cruzado: Un término cx_ix_j en una forma cuadrática, con $i \neq j$.

transformación (o **función**, o **mapeo**) ***T* de $\mathbb{R}^n$ a $\mathbb{R}^m$:** Regla que asigna a cada $\mathbf{x}$ en $\mathbb{R}^n$ un vector único $T(\mathbf{x})$ en $\mathbb{R}^m$. Notación: $T : \mathbb{R}^n \to \mathbb{R}^m$. También $T : V \to W$ denota una regla que asigna a cada $\mathbf{x}$ en V un vector único $T(\mathbf{x})$ en W.

transformación afín: Una función $T : \mathbb{R}^n \to \mathbb{R}^m$ de la forma $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, con A como una matriz de $m \times n$ y $\mathbf{b}$ en $\mathbb{R}^m$.

transformación de similitud: Transformación que cambia A en $P^{-1}AP$.

transformación lineal invertible: Transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ tal que existe una función $S : \mathbb{R}^n \to \mathbb{R}^n$ que satisface tanto a $T(S(\mathbf{x})) = \mathbf{x}$ como a $S(T(\mathbf{x})) = \mathbf{x}$ para todas las $\mathbf{x}$ en $\mathbb{R}^n$.

transformación lineal *T* (de un espacio vectorial V a un espacio vectorial W): Regla T que asigna a cada vector $\mathbf{x}$ en V un vector único $T(\mathbf{x})$ en W, tal que (*i*) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para toda $\mathbf{u}$, $\mathbf{v}$ en V, y (ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ para toda $\mathbf{u}$ en $\mathbf{V}$ y todos los escalares c. Notación: $T : V \to W$; también, $\mathbf{x} \mapsto A\mathbf{x}$ cuando $T : \mathbb{R}^n \to \mathbb{R}^m$ y A es la matriz estándar para T.

transformación matricial: Proyección $\mathbf{x} \mapsto A\mathbf{x}$, donde A es una matriz $m \times n$ y $\mathbf{x}$ representa cualquier vector en $\mathbb{R}^n$.

transpuesta (de A): Matriz A^T de $n \,\Delta\, m$ cuyas columnas son las filas correspondientes de la matriz A $m \times n$.

traslación (mediante un vector $\mathbf{p}$): Es la operación de sumar $\mathbf{p}$ a un vector o a cada vector en un conjunto dado.

trayectoria: Gráfica de una solución $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ de un sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, conectada con frecuencia mediante una curva delgada para lograr que la trayectoria resulte más fácilmente observable. También, la gráfica de $\mathbf{x}(t)$ para $t \geq 0$, cuando $\mathbf{x}(t)$ es una solución de una ecuación diferencial $\mathbf{x}'(t) = A\mathbf{x}(t)$.

traza (de una matriz cuadrada A): Suma de las entradas diagonales en A, denotada mediante tr A.

triangular superior en bloque (matriz): Es una matriz partida $A = [A_{ij}]$, tal que cada bloque A_{ij} es una matriz cero para $i > j$.

U

uno a uno (función): Función $T : \mathbb{R}^n \to \mathbb{R}^m$ tal que cada $\mathbf{b}$ en R^m es la imagen de *a lo más* una $\mathbf{x}$ en $\mathbb{R}^n$.

V

valor propio (de A): Escalar λ tal que la ecuación $A\mathbf{x} = \lambda\mathbf{x}$ tiene una solución para algún vector $\mathbf{x}$ distinto de cero.

valor propio complejo: Raíz no real de la ecuación característica de una matriz $n \times n$.

valor propio estrictamente dominante: Valor propio λ_1 de una matriz A con la propiedad de que $|\lambda_1| > |\lambda_k|$ para todos los demás valores propios λ_k de A.

valores singulares (de A): Son las raíces cuadradas positivas de los valores propiosde A^TA, establecidos en orden de magnitud descendente.

variable básica: En un sistema lineal, es una variable que corresponde a una columna pivote en la matriz de coeficientes.

variable libre: En un sistema lineal, cualquier variable que no es básica.

variables no correlacionadas: Son cualesquiera dos variables x_i y x_j (con $i \neq j$) que se encuentran sobre las coordenadas i-ésima y j-ésima de los vectores de observación en una matriz de observaciones, en forma tal que la covarianza s_{ij} es cero.

varianza (de una variable x_j): Es la entrada diagonal s_{jj} en la matriz de covarianza S para una matriz de observaciones, donde x_j varía sobre la j-ésima coordenada de los vectores de observación.

varianza total: Es la traza de la matriz de covarianza S de una matriz de observaciones.

vector cero: Es el vector único, denotado mediante $\mathbf{0}$, tal que $\mathbf{u} + \mathbf{0} = \mathbf{u}$ para toda $\mathbf{u}$. En $\mathbb{R}^n$, $\mathbf{0}$ es el vector cuyas entradas son todas iguales a cero.

vector columna: Matriz con sólo una columna, o una sola columna de una matriz que tiene varias columnas.

vector de consumo unitario: En el modelo de entrada y salida de Leontief, es un vector columna que enlista las entradas que un sector necesita para cada unidad de su salida; una columna de la matriz de consumo.

vector de coordenadas de x relativo a $\mathcal{B}$: El vector $[\mathbf{x}]_{\mathcal{B}}$ cuyas entradas son las coordenadas de $\mathbf{x}$ relativas a la base $\mathcal{B}$.

vector de demanda final (o **lista de demandas finales**): En el modelo de entrada y salida de Leontief, es el vector $\mathbf{d}$ que enlista el valor monetario de los bienes y servicios demandados por los diferentes sectores ubicados en la parte no productiva de la economía. El vector $\mathbf{d}$ puede representar la demanda del consumidor, el consumo del gobierno, el superávit de producción, las exportaciones, u otras demandas externas.

vector de equilibrio: *Vea* vector de estado estable.

vector de estado estable (para una matriz estocástica P): Vector de probabilidad $\mathbf{q}$ tal que $P\mathbf{q} = \mathbf{q}$.

vector de estado: Vector de probabilidad. En general, es un vector que describe el "estado" de un sistema físico, a menudo en conexión con una ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

vector de observación: Es el vector $\mathbf{y}$ en el modelo lineal $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde las entradas en $\mathbf{y}$ son los valores observados de una variable dependiente.

vector de probabilidad: Vector en $\mathbb{R}^n$ cuyas entradas son no negativas y suman 1.

vector de producción: En el modelo de entrada y salida de Leontief, es el vector que enlista las cantidades que serán producidas por los diferentes sectores de una economía.

vector fila: Matriz con sólo una fila, o una sola fila de una matriz que tiene varias filas.

vector parámetro: Vector desconocido $\boldsymbol{\beta}$ en el modelo lineal $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$.

vector propio (de A): Vector $\mathbf{x}$ *distinto de cero* tal que $AA\mathbf{x} = \lambda\mathbf{x}$ para algún escalar λ.

vector propio complejo: Vector $\mathbf{x}$ distinto de cero en $\mathbb{C}^n$ tal que $A\mathbf{x} = \lambda\mathbf{x}$, donde A es una matriz $n \times n$ y λ un valor propio complejo.

vector residual: Es la cantidad $\boldsymbol{\Delta}$ que aparece en el modelo lineal general: $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$; esto es, $\boldsymbol{\epsilon} = \mathbf{y} - X\boldsymbol{\beta}$, la diferencia entre los valores (de y) observados y los pronosticados.

vector unitario: Vector $\mathbf{v}$ tal que $\|\mathbf{v}\| = 1$.

vector: Lista de números; matriz con una sola columna. En general, cualquier elemento de un espacio vectorial.

vectores iguales: Vectores en $\mathbb{R}^n$ cuyas entradas correspondientes son las mismas.

vectores singulares derechos (de A): Son las columnas de V en la descomposición del valor singular $A = U\Sigma V^T$.

vectores singulares izquierdos (de A): En la descomposición en valores singulares $A = U\Sigma V^T$, son las columnas de U.

Respuestas a los ejercicios impares

CAPÍTULO 1

Sección 1.1, página 11

1 La solución es $(x_1, x_2) = (-8, 3)$, o simplemente $(-8, 3)$.

3. $(4/7, 9/7)$

5. Reemplace Fila2 por su suma con tres veces Fila3, y después sustituya Fila1 por su suma con -5 veces Fila3.

7. El conjunto solución está vacío.

9. $(4, 8, 5, 2)$ **11.** Inconsistente

13. $(5, 3, -1)$ **15.** Consistente

17. Las tres líneas tienen un punto en común.

19. $h \neq 2$ **21.** Todo h

23. Señale un enunciado como verdadero sólo cuando lo sea *siempre*. Incluir aquí las respuestas anularía el propósito de las preguntas de verdadero-falso, el cual es ayudar a aprender a leer cuidadosamente el texto. La *Guía de estudio* (Study Guide) le dirá dónde buscar las respuestas, pero el lector no debe consultar ninguna otra fuente sino hasta haber hecho un intento honesto de encontrar las respuestas por sí mismo.

25. $k + 2g + h = 0$

27. La reducción por filas de $\begin{bmatrix} 1 & 3 & f \\ c & d & g \end{bmatrix}$ a $\begin{bmatrix} 1 & 3 & f \\ 0 & d-3c & g-cf \end{bmatrix}$ muestra que $d - 3c$ debe ser diferente de cero, puesto que f y g son arbitrarias. De otra manera, para algunas selecciones de f y g, la segunda fila podría corresponder a una ecuación de la forma $0 = b$, donde b es distinta de cero. Así que $d \neq 3c$.

29. Intercambie Fila1 y Fila2; intercambie Fila1 y Fila2.

31. Reemplace Fila3 por Fila3 + (-4)Fila1; sustituya Fila3 por Fila3 + (4)Fila1.

33.
$$\begin{aligned} 4T_1 - T_2 - T_4 &= 30 \\ -T_1 + 4T_2 - T_3 &= 60 \\ -T_2 + 4T_3 - T_4 &= 70 \\ -T_1 - T_3 + 4T_4 &= 40 \end{aligned}$$

Sección 1.2, página 25

1. Forma escalonada reducida: a y b. Forma escalonada: d. No escalonada: c.

3. $\begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Cols pivote 1 y 2: $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 7 \\ 6 & 7 & 8 & 9 \end{bmatrix}$.

5. $\begin{bmatrix} \blacksquare & * \\ 0 & \blacksquare \end{bmatrix}$, $\begin{bmatrix} \blacksquare & * \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & \blacksquare \\ 0 & 0 \end{bmatrix}$

7. $\begin{cases} x_1 = -5 - 3x_2 \\ x_2 \text{ es libre} \\ x_3 = 3 \end{cases}$ **9.** $\begin{cases} x_1 = 4 + 5x_3 \\ x_2 = 5 + 6x_3 \\ x_3 \text{ es libre} \end{cases}$

11. $\begin{cases} x_1 = \frac{4}{3}x_2 - \frac{2}{3}x_3 \\ x_2 \text{ es libre} \\ x_3 \text{ es libre} \end{cases}$

13. $\begin{cases} x_1 = 5 + 3x_5 \\ x_2 = 1 + 4x_5 \\ x_3 \text{ es libre} \\ x_4 = 4 - 9x_5 \\ x_5 \text{ es libre} \end{cases}$

Nota: En la *Guía de estudio* (Study Guide) se analiza el error común de $x_3 = 0$.

15. **a.** Consistente, con una solución única

b. Inconsistente

17. $h = 7/2$

19. **a.** inconsistente cuando $h = 2$ y $k \neq 8$

b. Una solución única cuando $h \neq 2$

c. Muchas soluciones cuando $h = 2$ y $k = 8$

21. Lea cuidadosamente el texto y escriba las respuestas antes de consultar la *Guía de estudio* (Study Guide). Recuerde que un enunciado se considera verdadero solamente cuando lo es en todos los casos.

23. Sí. El sistema es consistente porque, con tres pivotes, debe haber un pivote en la tercera fila (la de abajo) de la matriz de coeficientes. La forma escalonada reducida no puede contener una fila de la forma [0 0 0 0 0 1].

25. Si la matriz de coeficientes tiene una posición pivote en cada fila, entonces hay una posición pivote en la fila de abajo, y no hay lugar para un pivote en la columna aumentada. Así que el sistema es consistente, de acuerdo con el teorema 2.

27. Si un sistema lineal es consistente, entonces la solución es única si, y sólo si, *cada columna en la matriz de coeficientes es una columna pivote; de otra forma, existe una cantidad infinita de soluciones*.

29. Un sistema subdeterminado siempre tiene más variables que ecuaciones. No puede haber más variables básicas que ecuaciones, así que debe existir por lo menos una variable libre. Es posible asignar a dicha variable un número infinito de valores diferentes. Si el sistema es consistente, cada valor diferente de una variable libre producirá una solución distinta.

31. Sí, un sistema de ecuaciones lineales con más ecuaciones que incógnitas puede ser consistente. El siguiente sistema tiene una solución ($x_1 = x_2 = 1$):

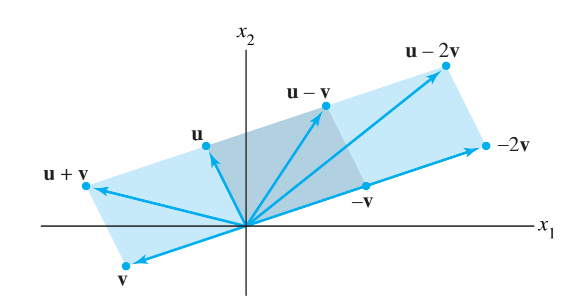
$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 0 \\ 3x_1 + 2x_2 &= 5 \end{aligned}$$

33. **[M]** $p(t) = 7 + 6t - t^2$

Sección 1.3, página 37

1. $\begin{bmatrix} -4 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}$

3.



5. $x_1 \begin{bmatrix} 6 \\ -1 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ -5 \end{bmatrix}$,

$$\begin{bmatrix} 6x_1 \\ -x_1 \\ 5x_1 \end{bmatrix} + \begin{bmatrix} -3x_2 \\ 4x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ -5 \end{bmatrix}, \begin{bmatrix} 6x_1 - 3x_2 \\ -x_1 + 4x_2 \\ 5x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ -5 \end{bmatrix}$$

$$\begin{aligned} 6x_1 - 3x_2 &= 1 \\ -x_1 + 4x_2 &= -7 \\ 5x_1 &= -5 \end{aligned}$$

Por lo general, los pasos intermedios no se muestran.

7. $\mathbf{a} = \mathbf{u} - 2\mathbf{v}$, $\mathbf{b} = 2\mathbf{u} - 2\mathbf{v}$, $\mathbf{c} = 2\mathbf{u} - 3.5\mathbf{v}$, $\mathbf{d} = 3\mathbf{u} - 4\mathbf{v}$

9. $x_1 \begin{bmatrix} 0 \\ 4 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ -1 \\ -8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

11. Sí, **b** es una combinación lineal de $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$.

13. No, **b** *no* es una combinación lineal de las columnas de A.

15. Por supuesto, se aceptan pesos no enteros, pero algunas selecciones sencillas son $0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 = \mathbf{0}$, y

$$1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 = \begin{bmatrix} 7 \\ 1 \\ -6 \end{bmatrix}, 0 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 = \begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix}$$

$$1 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 = \begin{bmatrix} 2 \\ 4 \\ -6 \end{bmatrix}, 1 \cdot \mathbf{v}_1 - 1 \cdot \mathbf{v}_2 = \begin{bmatrix} 12 \\ -2 \\ -6 \end{bmatrix}$$

17. $h = -17$

19. Gen$\{\mathbf{v}_1, \mathbf{v}_2\}$ es el conjunto de puntos sobre la línea que pasa por $\mathbf{v}_1$ y $\mathbf{0}$.

21. *Sugerencia:* Muestre que $\begin{bmatrix} 2 & 2 & h \\ -1 & 1 & k \end{bmatrix}$ es consistente para todas h y k. Explique lo que enseña este cálculo acerca de Gen{**u**, **v**}.

23. Antes de consultar la *Guía de estudio* (Study Guide) lea de manera cuidadosa toda la sección. Preste atención especial a las definiciones y los enunciados de los teoremas, y considere cualquier observación que esté antes o después de éstos.

25. **a.** No, a tres **b.** Sí, un número infinito
c. $\mathbf{a}_1 = 1 \cdot \mathbf{a}_1 + 0 \cdot \mathbf{a}_2 + 0 \cdot \mathbf{a}_3$

27. **a.** $5\mathbf{v}_1$ es la producción de 5 días de operación de la mina 1.
b. La producción total es $x_1\mathbf{v}_1 + x_2\mathbf{v}_2$, así que x_1 y x_2 deben satisfacer $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \begin{bmatrix} 150 \\ 2825 \end{bmatrix}$.
c. **[M]** 1.5 días para la mina 1 y 4 días para la mina 2

29. (1.3, .9, 0)

31. **a.** $\begin{bmatrix} 10/3 \\ 2 \end{bmatrix}$
b. Sume 3.5 g en (0, 1), sume .5 g en (8, 1), y sume 2 g en (2, 4).

33. Revise el problema de práctica 1 y después *escriba* una solución. La *Guía de estudio* (Study Guide) contiene una solución.

Sección 1.4, página 47

1. El producto no está definido porque el número de columnas (2) en la matriz de 3 × 2 no coincide con el número de entradas (3) en el vector.

3. $A\mathbf{x} = \begin{bmatrix} 6 & 5 \\ -4 & -3 \\ 7 & 6 \end{bmatrix}\begin{bmatrix} 2 \\ -3 \end{bmatrix} = 2 \cdot \begin{bmatrix} 6 \\ -4 \\ 7 \end{bmatrix} - 3 \cdot \begin{bmatrix} 5 \\ -3 \\ 6 \end{bmatrix}$

$= \begin{bmatrix} 12 \\ -8 \\ 14 \end{bmatrix} + \begin{bmatrix} -15 \\ 9 \\ -18 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ -4 \end{bmatrix}$, y

$A\mathbf{x} = \begin{bmatrix} 6 & 5 \\ -4 & -3 \\ 7 & 6 \end{bmatrix}\begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \cdot 2 + 5 \cdot (-3) \\ (-4) \cdot 2 + (-3) \cdot (-3) \\ 7 \cdot 2 + 6 \cdot (-3) \end{bmatrix}$

$= \begin{bmatrix} -3 \\ 1 \\ -4 \end{bmatrix}$. Muestre su trabajo aquí y para los ejercicios del 4 al 6, pero de ahí en adelante realice los cálculos mentalmente.

5. $5 \cdot \begin{bmatrix} 5 \\ -2 \end{bmatrix} - 1 \cdot \begin{bmatrix} 1 \\ -7 \end{bmatrix} + 3 \cdot \begin{bmatrix} -8 \\ 3 \end{bmatrix} - 2 \cdot \begin{bmatrix} 4 \\ -5 \end{bmatrix} = \begin{bmatrix} -8 \\ 16 \end{bmatrix}$

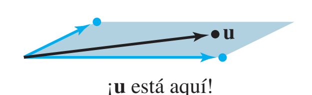
7. $\begin{bmatrix} 4 & -5 & 7 \\ -1 & 3 & -8 \\ 7 & -5 & 0 \\ -4 & 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -8 \\ 0 \\ -7 \end{bmatrix}$

9. $x_1 \begin{bmatrix} 3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} -5 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$ y

$\begin{bmatrix} 3 & 1 & -5 \\ 0 & 1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 2 & 4 & -2 \\ 0 & 1 & 5 & 2 \\ -2 & -4 & -3 & 9 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}$

13. Sí. (Justifique su respuesta.)



¡**u** está aquí!

15. La ecuación $A\mathbf{x} = \mathbf{b}$ no es consistente cuando $3b_1 + b_2$ es diferente de cero. (Muestre su trabajo.) El conjunto de **b** para el que la ecuación *es* consistente resulta ser una línea a través del origen —el conjunto de todos los puntos (b_1, b_2) que satisfacen $b_2 = -3b_1$.

17. Sólo tres filas contienen una posición pivote. La ecuación $A\mathbf{x} = \mathbf{b}$ *no* tiene una solución para cada **b** en $\mathbb{R}^4$, de acuerdo con el teorema 4.

19. El trabajo en el ejercicio 17 muestra que el enunciado (d) incluido en el teorema (4) es falso. Por lo tanto, los cuatro enunciados del teorema 4 son falsos. Así que no todos los vectores en $\mathbb{R}^4$ pueden escribirse como una combinación lineal de las columnas de A. Además, las columnas de A *no* generan $\mathbb{R}^4$.

21. La matriz $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$ no tiene un pivote en cada fila, así que las columnas de la matriz no generan $\mathbb{R}^4$, según el teorema 4. Esto es, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ no genera $\mathbb{R}^4$.

23. Lea el texto cuidadosamente y trate de señalar cada enunciado del ejercicio como verdadero o falso antes de consultar la *Guía de estudio* (Study Guide). Varias partes de los ejercicios 29 y 30 son *implicaciones* de la forma

"Si ⟨enunciado 1⟩, entonces ⟨enunciado 2⟩"

o, de manera equivalente,

"⟨enunciado 2⟩, si ⟨enunciado 1⟩"

Señale una implicación de este tipo como verdadera si (enunciado 2) es verdadero en todos los casos en que (enunciado 1) es verdadero.

25. $c_1 = -3$, $c_2 = -1$, $c_3 = 2$

27. $Q\mathbf{x} = \mathbf{v}$, donde $Q = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3]$ y $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Nota: Si su respuesta es la ecuación $A\mathbf{x} = \mathbf{b}$, debe especificar de cuáles A y $\mathbf{b}$ se trata.

29. *Sugerencia:* Inicie con cualquier matriz B de 3×3 en forma escalonada que tenga tres posiciones pivote.

31. *Escriba* su solución antes de consultar la *Guía de estudio* (*Study Guide*).

33. *Pista:* ¿Cuántas columnas pivote tiene A? ¿Por qué?

35. Dadas $A\mathbf{x}_1 = \mathbf{y}_1$ y $A\mathbf{x}_2 = \mathbf{y}_2$, se pide mostrar que la ecuación $A\mathbf{x} = \mathbf{w}$ tiene una solución, donde $\mathbf{w} = \mathbf{y}_1 + \mathbf{y}_2$. Observe que $\mathbf{w} = A\mathbf{x}_1 + A\mathbf{x}_2$ y utilice el teorema 5(a) con $\mathbf{x}_1$ y $\mathbf{x}_2$ en lugar de $\mathbf{u}$ y $\mathbf{v}$, respectivamente. Esto es, $\mathbf{w} = A\mathbf{x}_1 + A\mathbf{x}_2 = A(\mathbf{x}_1 + \mathbf{x}_2)$. De manera que el vector $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ es una solución de $\mathbf{w} = A\mathbf{x}$.

37. **[M]** Las columnas no generan $\mathbb{R}^4$.

39. **[M]** Las columnas generan $\mathbb{R}^4$.

41. **[M]** Borre la columna 4 de la matriz del ejercicio 39. También es posible borrar la columna 3 en lugar de la 4.

Sección 1.5, página 55

1. El sistema tiene una solución no trivial porque existe una variable libre, x_3.

3. El sistema tiene una solución no trivial porque existe una variable libre, x_3.

5. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$

7. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -9 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix}$

9. $\mathbf{x} = x_2 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$

11. *Sugerencia:* El sistema que se deriva de la forma escalonada *reducida* es

$$\begin{aligned} x_1 - 4x_2 \qquad\quad + 5x_6 &= 0 \\ x_3 \quad - \ x_6 &= 0 \\ x_5 - 4x_6 &= 0 \\ 0 &= 0 \end{aligned}$$

Las variables básicas son x_1, x_3 y x_5. Las variables restantes son libres. En la *Guía de estudio* (*Study Guide*) se analizan dos errores que se cometen en este tipo de problemas.

13. $\mathbf{x} = \begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ -7 \\ 1 \end{bmatrix} = \mathbf{p} + x_3\mathbf{q}$. Geométricamente, el conjunto solución es la línea que pasa por $\begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix}$ y es paralela a $\begin{bmatrix} 4 \\ -7 \\ 1 \end{bmatrix}$.

15. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$. El conjunto solución es la línea que pasa por $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$, paralela a la línea constituida como conjunto solución del sistema homogéneo del ejercicio 5.

17. Sean $\mathbf{u} = \begin{bmatrix} -9 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}$. La solución de la ecuación homogénea es $\mathbf{x} = x_2\mathbf{u} + x_3\mathbf{v}$, el plano que pasa por el origen generado mediante $\mathbf{u}$ y $\mathbf{v}$. El conjunto solución del sistema no homogéneo es $\mathbf{x} = \mathbf{p} + x_2\mathbf{u} + x_3\mathbf{v}$, el plano que pasa por $\mathbf{p}$ paralelo al conjunto solución de la ecuación homogénea.

19. $\mathbf{x} = \mathbf{a} + t\mathbf{b}$, donde t representa un parámetro, o $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} -5 \\ 3 \end{bmatrix}$, o bien $\begin{cases} x_1 = -2 - 5t \\ x_2 = 3t \end{cases}$

21. $\mathbf{x} = \mathbf{p} + t(\mathbf{q} - \mathbf{p}) = \begin{bmatrix} 2 \\ -5 \end{bmatrix} + t \begin{bmatrix} -5 \\ 6 \end{bmatrix}$

23. Es importante leer el texto de manera cuidadosa y escribir sus respuestas. Después de esto, consulte la *Guía de estudio* (*Study Guide*), de ser necesario.

25. **a.** $A\mathbf{w} = A(\mathbf{p} + \mathbf{v}_h) = A\mathbf{p} + A\mathbf{v}_h = \mathbf{b} + \mathbf{0} = \mathbf{b}$
b. $A\mathbf{v}_h = A(\mathbf{w} - \mathbf{p}) = A\mathbf{w} - A\mathbf{p} = \mathbf{b} - \mathbf{b} = \mathbf{0}$

27. Cuando A es la matriz cero de 3×3, *todo* $\mathbf{x}$ en $\mathbb{R}^3$ satisface $A\mathbf{x} = \mathbf{0}$. De modo que el conjunto solución consiste en todos los vectores en $\mathbb{R}^3$.

29. **a.** Cuando A es una matriz de 3×3 con tres posiciones pivote, la ecuación $A\mathbf{x} = \mathbf{0}$ no tiene variables libres y, por lo tanto, no tiene solución no trivial.

b. Con tres posiciones pivote, A tiene una posición pivote en cada una de sus tres filas. De acuerdo con el teorema 4 de la sección 1.4, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cualquier $\mathbf{b}$ posible. En el ejercicio la palabra "posible" significa que los únicos vectores considerados en este caso son los de $\mathbb{R}^3$, porque A tiene tres filas.

31. **a.** Cuando A es una matriz de 3×2 con dos posiciones pivote, cada columna es una columna pivote. Entonces la ecuación $A\mathbf{x} = \mathbf{0}$ no tiene variables libres y, por lo tanto, tampoco tiene solución no trivial.

b. Con dos posiciones pivote y tres filas, A no puede tener un pivote en cada fila. Por lo tanto, la ecuación $A\mathbf{x} = \mathbf{b}$

no puede tener una solución para todo **b** posible (en $\mathbb{R}^3$), de acuerdo con el teorema 4 de la sección 1.4.

33. Una respuesta: $\mathbf{x} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$

35. Su ejemplo debe tener la propiedad de que la suma de las entradas en cada fila es cero. ¿Por qué?

37. Una respuesta es $A = \begin{bmatrix} 1 & -4 \\ 1 & -4 \end{bmatrix}$. En la *Guía de estudio* (*Study Guide*) se muestra cómo analizar el problema con el fin de construir A. Si **b** es cualquier vectorque *no* sea un múltiplo de la primera columna de A, entonces el conjunto solución de $A\mathbf{x} = \mathbf{b}$ está vacío y, por lo tanto, no puede formarse al trasladar el conjunto solución de $A\mathbf{x} = \mathbf{b}$. Esto no contradice al teorema 6, porque dicho teorema se aplica cuando la ecuación $A\mathbf{x} = \mathbf{b}$ tiene un conjunto solución no vacío.

39. Si c es un escalar, entonces $A(c\mathbf{u}) = cA\mathbf{u}$, según el teorema 5(b) de la sección 1.4. Si **u** satisface $A\mathbf{x} = \mathbf{0}$, entonces $A\mathbf{u} = \mathbf{0}$, $cA\mathbf{u} = c \cdot \mathbf{0} = \mathbf{0}$, entonces $A(c\mathbf{u}) = \mathbf{0}$.

Sección 1.6, página 63

1. La solución general es $p_{\text{Bienes}} = .875 p_{\text{Servicios}}$, con $p_{\text{Servicios}}$ libre. Una solución de equilibrio es $p_{\text{Servicios}} = 1000$ y $p_{\text{Bienes}} = 875$. Usando fracciones, la solución general puede escribirse como $p_{\text{Bienes}} = (7/8) p_{\text{Servicios}}$, y una selección natural de precios podría ser $p_{\text{Servicios}} = 80$ y $p_{\text{Bienes}} = 70$. Sólo es importante la *razón* de los precios. El equilibrio económico no se ve afectado por un cambio proporcional en los precios.

3. a.

	Distribución de la producción de:				
	QyM	CyE	Maq.		
Salida	↓	↓	↓	Entrada	Comprada por:
	.2	.8	.4	→	QyM
	.3	.1	.4	→	CyE
	.5	.1	.2	→	Maq.

b. $\begin{bmatrix} .8 & -.8 & -.4 & 0 \\ -.3 & .9 & -.4 & 0 \\ -.5 & -.1 & .8 & 0 \end{bmatrix}$

c. **[M]** $p_{\text{Químicos}} = 141.7$, $p_{\text{Combustibles}} = 91.7$, $p_{\text{Maquinaria}} = 100$. Hasta dos cifras significativas, $p_{\text{Químicos}} = 140$, $p_{\text{Combustibles}} = 92$, $p_{\text{Maquinaria}} = 100$.

5. $B_2S_3 + 6H_2O \rightarrow 2H_3BO_3 + 3H_2S$

7. $3NaHCO_3 + H_3C_6H_5O_7 \rightarrow Na_3C_6H_5O_7 + 3H_2O + 3CO_2$

9. **[M]** $15PbN_6 + 44CrMn_2O_8 \rightarrow$
$5Pb_3O_4 + 22Cr_2O_3 + 88MnO_2 + 90NO$

11. $\begin{cases} x_1 = 20 - x_3 \\ x_2 = 60 + x_3 \\ x_3 \text{ es libre} \\ x_4 = 60 \end{cases}$ El mayor valor de x_3 es 20.

13. a. $\begin{cases} x_1 = x_3 - 40 \\ x_2 = x_3 + 10 \\ x_3 \text{ es libre} \\ x_4 = x_6 + 50 \\ x_5 = x_6 + 60 \\ x_6 \text{ es libre} \end{cases}$ **b.** $\begin{cases} x_2 = 50 \\ x_3 = 40 \\ x_4 = 50 \\ x_5 = 60 \end{cases}$

Sección 1.7, página 71

Justifique sus respuestas a los ejercicios 1 a 22.

1. Linealmente independiente **3.** Linealmente dependiente

5. Linealmente independiente **7.** Linealmente dependiente

9. a. Ninguna h **b.** Toda h

11. $h = 6$ **13.** Toda h

15. Linealmente dependiente **17.** Linealmente dependiente

19. Linealmente dependiente

21. Si consulta la *Guía de estudio* (*Study Guide*) antes de hacer un buen esfuerzo por responder las preguntas de verdadero o falso destruirá la mayor parte de su valor.

23. $\begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \end{bmatrix}$ **25.** $\begin{bmatrix} \blacksquare & * \\ 0 & \blacksquare \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ y $\begin{bmatrix} 0 & \blacksquare \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

27. Las cinco columnas de la matriz A de 7×5 deben ser columnas pivote. De otra forma, la ecuación $A\mathbf{x} = \mathbf{0}$ tendría una variable libre, en cuyo caso las columnas de A serían linealmente dependientes.

29. A: Cualquier matriz de 3×2 con dos columnas distintas de cero tales que ninguna de las columnas sea múltiplo de la otra. En este caso, las columnas son linealmente independientes, y, por lo tanto, la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.
B: Cualquier matriz de 3×2 donde una columna sea múltiplo de la otra.

31. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

33. Verdadero, según el teorema 7. En la *Guía de estudio* (*Study Guide*) se proporciona otra justificación.

35. Falso. El vector $\mathbf{v}_1$ podría ser el vector cero.

37. Verdadero. Una relación de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ puede ampliarse a una relación de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$ al colocar un peso cero en $\mathbf{v}_4$.

39. El lector debe ser capaz de trabajar con este importante problema sin ayuda. *Escriba* su solución antes de consultar la *Guía de estudio* (*Study Guide*) .

41. **[M]** $B = \begin{bmatrix} 8 & -3 & 2 \\ -9 & 4 & -7 \\ 6 & -2 & 4 \\ 5 & -1 & 10 \end{bmatrix}$. También son posibles otras selecciones.

43. **[M]** Cada columna de A que no sea una columna de B está en el conjunto generado por las columnas de B.

Sección 1.8, página 79

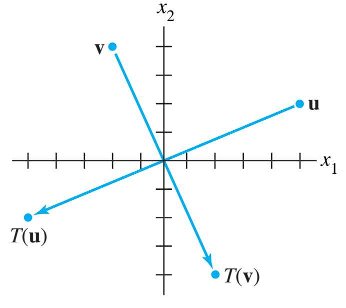
1. $\begin{bmatrix} 2 \\ -6 \end{bmatrix}$, $\begin{bmatrix} 2a \\ 2b \end{bmatrix}$ **3.** $\mathbf{x} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$, solución única

5. $\mathbf{x} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$, no es única **7.** $a = 5, b = 6$

9. $\mathbf{x} = x_3 \begin{bmatrix} 9 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -7 \\ -3 \\ 0 \\ 1 \end{bmatrix}$

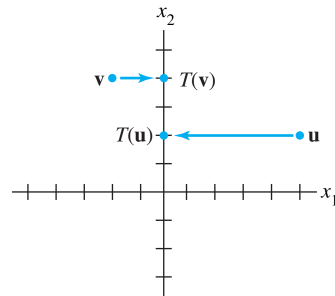
11. Sí, porque el sistema representado por $[A \quad \mathbf{b}]$ es consistente.

13.



Una reflexión a través del origen

15.



Una proyección sobre el eje x_2.

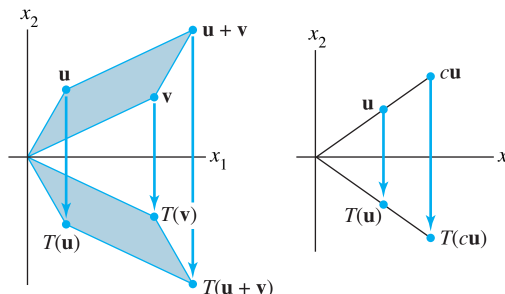
17. $\begin{bmatrix} 6 \\ 3 \end{bmatrix}$, $\begin{bmatrix} -2 \\ 6 \end{bmatrix}$, $\begin{bmatrix} 4 \\ 9 \end{bmatrix}$ **19.** $\begin{bmatrix} 13 \\ 7 \end{bmatrix}$, $\begin{bmatrix} 2x_1 - x_2 \\ 5x_1 + 6x_2 \end{bmatrix}$

21. Lea el texto de manera cuidadosa y escriba sus respuestas antes de consultar la *Guía de estudio* (*Study Guide*). Observe que el ejercicio 21(e) es una afirmación de la forma

"⟨enunciado 1⟩ si, y sólo si, ⟨enunciado 2⟩"

Señale una afirmación de este tipo como verdadera si ⟨enunciado 1⟩ es verdadero siempre que ⟨enunciado 2⟩ sea verdadero, *y* también ⟨enunciado 2⟩ es verdadero siempre que ⟨enunciado 1⟩ sea verdadero.

23.



25. *Sugerencia:* Muestre que la imagen (esto es, el conjunto de imágenes de todos los puntos de una línea) pueden representarse mediante la ecuación paramétrica de una línea.

27. **a.** La línea que pasa por $\mathbf{p}$ y $\mathbf{q}$ es paralela a $\mathbf{q} - \mathbf{p}$. (Vea la figura 7 de la sección 1.5.) Como $\mathbf{p}$ está sobre la línea, la ecuación de la línea es $\mathbf{x} = \mathbf{p} + t(\mathbf{q} - \mathbf{p})$. Vuelva a escribir esto como $\mathbf{x} = \mathbf{p} - t\mathbf{p} + t\mathbf{q}$ y $\mathbf{x} = (1 - t)\mathbf{p} + t\mathbf{q}$.

b. Considere $\mathbf{x} = (1 - t)\mathbf{p} + t\mathbf{q}$ para t tal que $0 \le t \le 1$. Entonces, por la linealidad de T, para $0 \le t \le 1$

$$T(\mathbf{x}) = T((1 - t)\mathbf{p} + t\mathbf{q}) = (1 - t)T(\mathbf{p}) + tT(\mathbf{q}) \qquad *$$

Si $T(\mathbf{p})$ y $T(\mathbf{q})$ son distintos, entonces (*) es la ecuación para el segmento de línea entre $T(\mathbf{p})$ y $T(\mathbf{q})$, como se muestra en el inciso (a). De otra manera, el conjunto de

imágenes es sólo el punto $T(\mathbf{p})$, porque
$(1-t)T(\mathbf{p}) + tT(\mathbf{q}) = (1-t)T(\mathbf{p}) + tT(\mathbf{p}) = T(\mathbf{p})$

29. **a.** Cuando $b = 0$, $f(x) = mx$. En este caso, para toda x, y en $\mathbb{R}$ y todos los escalares c y d,

$$\begin{aligned} f(cx+dy) &= m(cx+dy) = mcx + mdy \\ &= c(mx) + d(my) = c \cdot f(x) + d \cdot f(y) \end{aligned}$$

Esto muestra que f es lineal.

b. Cuando $f(x) = mx + b$, con b diferente de cero, $f(0) = m(0) + b = b \neq 0$.

c. En cálculo, f se llama "función lineal" porque la gráfica de f es una línea.

31. *Sugerencia:* Como $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente, puede escribirse alguna ecuación y trabajar con ella.

33. Una posibilidad es mostrar que T no mapea el vector cero en el vector cero, algo que cualquier transformación lineal *sí* hace: $T(0, 0) = (0, 4, 0)$.

35. Tome $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^3$ y sean c y d escalares. Entonces

$$c\mathbf{u} + d\mathbf{v} = (cu_1 + dv_1, cu_2 + dv_2, cu_3 + dv_3)$$

La transformación T es lineal porque

$$\begin{aligned} T(c\mathbf{u} + d\mathbf{v}) &= (cu_1 + dv_1, cu_2 + dv_2, -(cu_3 + dv_3)) \\ &= (cu_1 + dv_1, cu_2 + dv_2, -cu_3 - dv_3) \\ &= (cu_1, cu_2, -cu_3) + (dv_1, dv_2, -dv_3) \\ &= c(u_1, u_2, -u_3) + d(v_1, v_2, -v_3) \\ &= cT(\mathbf{u}) + dT(\mathbf{v}) \end{aligned}$$

37. **[M]** Todos los múltiplos de $(7, 9, 0, 2)$

39. **[M]** Sí. Una opción para $\mathbf{x}$ es $(4, 7, 1, 0)$.

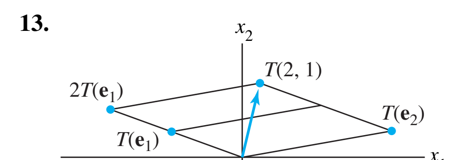
Sección 1.9, página 90

1. $\begin{bmatrix} 3 & -5 \\ 1 & 2 \\ 3 & 0 \\ 1 & 0 \end{bmatrix}$ **3.** $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ **5.** $\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$

7. $\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ **9.** $\begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}$

11. La transformación T descrita mapea $\mathbf{e}_1$ en $-\mathbf{e}_1$ y $\mathbf{e}_2$ en $-\mathbf{e}_2$. Una rotación a través de π radianes también mapea $\mathbf{e}_1$ en $-\mathbf{e}_1$ y $\mathbf{e}_2$ en $-\mathbf{e}_2$. Como una transformación lineal está completamente determinada por lo que le hace a las columnas de la matriz identidad, la transformación de rotación tiene el mismo efecto que T en cualquier vector de $\mathbb{R}^2$.

13.



15. $\begin{bmatrix} 3 & 0 & -2 \\ 4 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}$ **17.** $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

19. $\begin{bmatrix} 1 & -5 & 4 \\ 0 & 1 & -6 \end{bmatrix}$ **21.** $\mathbf{x} = \begin{bmatrix} 7 \\ -4 \end{bmatrix}$

23. Responda las preguntas antes de consultar la *Guía de estudio* (*Study Guide*).

Justifique sus respuestas a los ejercicios 25 a 28.

25. No es uno a uno y no mapea $\mathbb{R}^4$ sobre $\mathbb{R}^4$.

27. No es uno a uno pero mapea $\mathbb{R}^3$ sobre $\mathbb{R}^2$.

29. $\begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \\ 0 & 0 & 0 \end{bmatrix}$

31. n. (Explique por qué y después consulte la *Guía de estudio* (*Study Guide*).

33. *Sugerencia:* Si $\mathbf{e}_j$ es la j-ésima columna de I_n, entonces $B\mathbf{e}_j$ es la j-ésima columna de B.

35. *Sugerencia:* ¿Es posible que $m > n$? ¿Y puede ser que $m < n$?

37. **[M]** No. (Explique por qué.)

39. **[M]** No. (Explique por qué.)

Sección 1.10, página 99

1. **a.** $x_1 \begin{bmatrix} 110 \\ 4 \\ 20 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 130 \\ 3 \\ 18 \\ 5 \end{bmatrix} = \begin{bmatrix} 295 \\ 9 \\ 48 \\ 8 \end{bmatrix}$, donde x_1 es el número de porciones de Cheerios y x_2 es el número de porciones de Cereal 100% Natural.

b. $\begin{bmatrix} 110 & 130 \\ 4 & 3 \\ 20 & 18 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 295 \\ 9 \\ 48 \\ 8 \end{bmatrix}$. Mezcla de 1.5 porciones de Cheerios junto con 1 porción de Cereal 100% Natural.

3. a. $\begin{bmatrix} 36 & 51 & 13 & 80 \\ 52 & 34 & 74 & 0 \\ 0 & 7 & 1.1 & 3.4 \\ 1.26 & .19 & .8 & .18 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 33 \\ 45 \\ 3 \\ .8 \end{bmatrix}$, donde

donde $x_1, \ldots, x_4$ representan los números de unidades (100 g) de leche desgrasada, harina de soya, suero y proteína de soya aislada, respectivamente, que se usarán en la mezcla.

b. **[M]** La "solución" es $x_1 = .64$, $x_2 = .54$, $x_3 = -.09$, $x_4 = -.21$. Esta solución no es factible, porque la mezcla no puede incluir cantidades negativas de suero y proteína de soya aislada.

5. $R\mathbf{i} = \mathbf{v}$, $\begin{bmatrix} 5 & -2 & 0 & 0 \\ -2 & 11 & -3 & 0 \\ 0 & -3 & 17 & -4 \\ 0 & 0 & -4 & 25 \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 40 \\ -30 \\ 20 \\ -10 \end{bmatrix}$

[M]: $\mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 7.56 \\ -1.10 \\ .93 \\ -.25 \end{bmatrix}$

7. $R\mathbf{i} = \mathbf{v}$, $\begin{bmatrix} 12 & -7 & 0 & -4 \\ -7 & 15 & -6 & 0 \\ 0 & -6 & 14 & -5 \\ -4 & 0 & -5 & 13 \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 40 \\ 30 \\ 20 \\ -10 \end{bmatrix}$

[M]: $\mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 11.43 \\ 10.55 \\ 8.04 \\ 5.84 \end{bmatrix}$

9. $\mathbf{x}_{k+1} = M\mathbf{x}_k$ para $k = 0, 1, 2, \ldots,$ donde

$M = \begin{bmatrix} .95 & .04 \\ .05 & .96 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$

En 2002 (cuando $k = 2$) la población es $\mathbf{x}_2 = \begin{bmatrix} 573{,}260 \\ 426{,}740 \end{bmatrix}$.

11. a. $M = \begin{bmatrix} .98285 & .00258 \\ .01715 & .99742 \end{bmatrix}$

b. **[M]** $\mathbf{x}_{10} = \begin{bmatrix} 30{,}223{,}000 \\ 218{,}487{,}000 \end{bmatrix}$ Al millar más cercano

13. **[M] a.** La población de la ciudad disminuye. Después de 7 años, las poblaciones son aproximadamente iguales, pero la población de la ciudad continúa decayendo. Después de 20 años, sólo hay 417,000 personas en la ciudad. (*Nota:* 417,456 redondeado al millar más cercano.) Sin embargo, los cambios en la población parecen ser menores cada año.

b. La población de la ciudad aumenta lentamente. Después de 20 años, la población ha crecido de 350,000 a alrededor de 370,000 habitantes.

Capítulo 1, ejercicios suplementarios, página 102

1. **a.** F **b.** F **c.** T **d.** F **e.** T
f. T **g.** F **h.** F **i.** T **j.** F
k. T **l.** F **m.** T **n.** T **o.** T
p. T **q.** F **r.** T **s.** F **t.** F
u. F **v.** T **w.** T **x.** F **y.** T **z.** F

3. a. Cualquier sistema lineal consistente cuya forma escalonada sea

$\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$ o $\begin{bmatrix} \blacksquare & * & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$

o $\begin{bmatrix} 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$

b. Cualquier sistema lineal consistente cuya forma escalonada reducida sea I_3.

c. Cualquier sistema lineal inconsistente de tres ecuaciones en tres variables.

5. a. El conjunto solución: (i) está vacío si $h = 12$ y $k \neq 2$; (ii) contiene una solución única si $h \neq 12$; (iii) contiene un número infinito de soluciones si $h = 12$ y $k = 2$.

b. El conjunto solución está vacío si $k + 3h = 0$; de otro modo, el conjunto solución contiene una solución única.

7. a. Establezca $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -5 \\ 7 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ -5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}$, y

$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. "Determine si $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ generan $\mathbb{R}^3$."

Solución: No.

b. Establezca $A = \begin{bmatrix} 2 & -4 & -2 \\ -5 & 1 & 1 \\ 7 & -5 & -3 \end{bmatrix}$. "Determine si las columnas de A generan $\mathbb{R}^3$."

c. Defina $T(\mathbf{x}) = A\mathbf{x}$. "Determine si T mapea $\mathbb{R}^3$ sobre $\mathbb{R}^3$."

9. $\begin{bmatrix} 5 \\ 6 \end{bmatrix} = \frac{4}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{7}{3}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ o $\begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 8/3 \\ 4/3 \end{bmatrix} + \begin{bmatrix} 7/3 \\ 14/3 \end{bmatrix}$

Sugerencia: Estructure una "cuadrícula" sobre el plano x_1x_2 determinado por $\mathbf{a}_1$ y $\mathbf{a}_2$.

11. Un conjunto solución es una línea cuando el sistema tiene una variable libre. Si la matriz de coeficientes es de 2×3, entonces dos de las columnas deben ser columnas pivote. Por ejemplo, tome $\begin{bmatrix} 1 & 2 & * \\ 0 & 3 & * \end{bmatrix}$. Anote cualquier número en la columna 3. La matriz resultante estará en forma escalonada. Efectúe una operación de sustitución por filas sobre la segunda fila para crear una matriz que *no* esté en forma escalonada, tal como $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 1 & 5 & 2 \end{bmatrix}$.

12. *Sugerencia:* ¿Cuántas variables libres hay en la ecuación $A\mathbf{x} = \mathbf{0}$?

13. $E = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

15. a. Si los tres vectores son linealmente independientes, entonces a, c y f deben ser todos diferentes de cero.

b. Los números $a, \ldots, f$ pueden tener cualesquiera valores.

16. *Sugerencia:* Enliste las columnas de derecha a izquierda como $\mathbf{v}_1, \ldots, \mathbf{v}_4$.

17. *Sugerencia:* Use el teorema 7.

19. Sea M la línea que pasa por el origen y es paralela a la línea que pasa por $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Entonces $\mathbf{v}_2 - \mathbf{v}_1$ y $\mathbf{v}_3 - \mathbf{v}_1$ están ambos sobre M. Por lo tanto, uno de estos dos vectores es múltiplo del otro, por ejemplo, $\mathbf{v}_2 - \mathbf{v}_1 = k(\mathbf{v}_3 - \mathbf{v}_1)$. Esta ecuación produce una relación de dependencia lineal: $(k - 1)\mathbf{v}_1 + \mathbf{v}_2 - k\mathbf{v}_3 = \mathbf{0}$.

Una segunda solución: Una ecuación paramétrica de la línea es $\mathbf{x} = \mathbf{v}_1 + t(\mathbf{v}_2 - \mathbf{v}_1)$. Como $\mathbf{v}_3$ está sobre la línea, existe un t_0 tal que $\mathbf{v}_3 = \mathbf{v}_1 + t_0(\mathbf{v}_2 - \mathbf{v}_1) = (1 - t_0)\mathbf{v}_1 + t_0\mathbf{v}_2$. Por lo tanto, $\mathbf{v}_3$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$, y $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente.

21. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ **23.** $a = 4/5$ y $b = -3/5$

25. a. El vector enlista los números de departamentos de tres, dos y un dormitorios que hay cuando se construyen x_1 pisos con el plan A.

b. $x_1 \begin{bmatrix} 3 \\ 7 \\ 8 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 4 \\ 8 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ 3 \\ 9 \end{bmatrix}$

c. **[M]** Utilice 2 pisos del plan A y 15 pisos del plan B. O use 6 pisos del plan A, 2 pisos del plan B, y 8 pisos del plan C. Éstas son las únicas soluciones factibles. Hay otras soluciones matemáticas, pero requieren de un número negativo de pisos en uno o dos de los planes, lo cual no tiene sentido físico.

CAPÍTULO 2

Sección 2.1, página 116

1. $\begin{bmatrix} -4 & 0 & 2 \\ -8 & 10 & -4 \end{bmatrix}$, $\begin{bmatrix} 3 & -5 & 3 \\ -7 & 6 & -7 \end{bmatrix}$, no definida, $\begin{bmatrix} 1 & 13 \\ -7 & -6 \end{bmatrix}$

3. $\begin{bmatrix} -1 & 1 \\ -5 & 5 \end{bmatrix}$, $\begin{bmatrix} 12 & -3 \\ 15 & -6 \end{bmatrix}$

5. a. $A\mathbf{b}_1 = \begin{bmatrix} -7 \\ 7 \\ 12 \end{bmatrix}$, $A\mathbf{b}_2 = \begin{bmatrix} 4 \\ -6 \\ -7 \end{bmatrix}$, $AB = \begin{bmatrix} -7 & 4 \\ 7 & -6 \\ 12 & -7 \end{bmatrix}$

b. $AB = \begin{bmatrix} -1 \cdot 3 + 2(-2) & -1(-2) + 2 \cdot 1 \\ 5 \cdot 3 + 4(-2) & 5(-2) + 4 \cdot 1 \\ 2 \cdot 3 - 3(-2) & 2(-2) - 3 \cdot 1 \end{bmatrix} = \begin{bmatrix} -7 & 4 \\ 7 & -6 \\ 12 & -7 \end{bmatrix}$

7. 3×7 **9.** $k - 5$

11. $AD = \begin{bmatrix} 2 & 3 & 5 \\ 2 & 6 & 15 \\ 2 & 12 & 25 \end{bmatrix}$, $DA = \begin{bmatrix} 2 & 2 & 2 \\ 3 & 6 & 9 \\ 5 & 20 & 25 \end{bmatrix}$

La multiplicación derecha por D multiplica cada *columna* de A por la entrada diagonal correspondiente de D. La multiplicación izquierda por D multiplica cada *fila* por la entrada diagonal correspondiente de D. La *Guía de estudio* (*Study Guide*) dice cómo hacer $AB = BA$, pero es preferible que lo intente por sí mismo antes de verlo allí.

13. *Sugerencia:* Una de las dos matrices es Q.

15. Responda las preguntas antes de consultar la *Guía de estudio* (*Study Guide*).

17. $\mathbf{b}_1 = \begin{bmatrix} 7 \\ 4 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -8 \\ -5 \end{bmatrix}$

19. La tercera columna de AB es la suma de las primeras dos columnas de AB. He aquí por qué. Escriba $B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]$. Por definición, la tercera columna de AB es $A\mathbf{b}_3$. Si $\mathbf{b}_3 = \mathbf{b}_1 + \mathbf{b}_2$, entonces $A\mathbf{b}_3 = A(\mathbf{b}_1 + \mathbf{b}_2) = A\mathbf{b}_1 + A\mathbf{b}_2$, de acuerdo con una propiedad de la multiplicación matriz-vector.

21. Las columnas de A son linealmente dependientes. ¿Por qué?

23. *Sugerencia:* Suponga que $\mathbf{x}$ satisface $A\mathbf{x} = \mathbf{0}$, y muestre que $\mathbf{x}$ debe ser $\mathbf{0}$.

25. *Sugerencia:* Use los resultados de los ejercicios 23 y 24, y aplique la ley asociativa de la multiplicación al producto CAD.

27. $\mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u} = -2a + 3b - 4c$,

$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} -2a & -2b & -2c \\ 3a & 3b & 3c \\ -4a & -4b & -4c \end{bmatrix}$

$\mathbf{v}\mathbf{u}^T = \begin{bmatrix} -2a & 3a & -4a \\ -2b & 3b & -4b \\ -2c & 3c & -4c \end{bmatrix}$

29. *Sugerencia:* De acuerdo con el teorema 2(b), muestre que las entradas (i, j) de $A(B + C)$ y de $AB + AC$ son iguales.

31. *Sugerencia:* Use la definición del producto I_mA y el hecho de que $I_m\mathbf{x} = \mathbf{x}$ para $\mathbf{x}$ en $\mathbb{R}^m$.

33. *Sugerencia:* Primero escriba la entrada (i, j) de $(AB)^T$, que es la entrada (j, i) de AB. Luego, para calcular la entrada (i, j) en B^TA^T, aplique los hechos de que las entradas en la fila i de B^T son $b_{1i}, \ldots, b_{ni}$, porque provienen de la columna i de B, y que las entradas en la columna j de A^T son $a_{j1}, \ldots, a_{jn}$, porque provienen de la fila j de A.

35. **[M]** Aquí su respuesta dependerá de la selección del programa de matrices. Para MATLAB, utilice el comando `help` para leer acerca de `zeros, ones, eye` y `diag`. Para la TI-86, estudie las instrucciones `dim, fill` e `iden`. La TI-86 no tiene comando "diagonal".

37. **[M]** Presente sus resultados e informe acerca de sus conclusiones.

39. **[M]** La matriz S "cambia" las entradas en un vector (a, b, c, d, e) para producir $(b, c, d, e, 0)$. S^5 es la matriz cero de 5×5. Lo mismo pasa con S^6.

Sección 2.2, página 126

1. $\begin{bmatrix} 2 & -3 \\ -5/2 & 4 \end{bmatrix}$ **3.** $-\frac{1}{5}\begin{bmatrix} -5 & -5 \\ 7 & 8 \end{bmatrix}$ o $\begin{bmatrix} 1 & 1 \\ -7/5 & -8/5 \end{bmatrix}$

5. $x_1 = 7$ y $x_2 = -9$

7. **a** y **b**: $\begin{bmatrix} -9 \\ 4 \end{bmatrix}, \begin{bmatrix} 11 \\ -5 \end{bmatrix}, \begin{bmatrix} 6 \\ -2 \end{bmatrix}$, y $\begin{bmatrix} 13 \\ -5 \end{bmatrix}$

9. Escriba sus respuestas antes de consultar la *Guía de estudio* (*Study Guide*).

11. Se puede hacer la demostración tomando como modelo la del teorema 5.

13. $AB = AC \Rightarrow A^{-1}AB = A^{-1}AC \Rightarrow IB = IC \Rightarrow B = C$. No, en general, B y C pueden ser diferentes si A no es invertible. Vea el ejercicio 10 de la sección 2.1.

15. $D = C^{-1}B^{-1}A^{-1}$. Muestre que D funciona.

17. $A = BCB^{-1}$

19. Después de encontrar $X = CB - A$, muestre que X es una solución.

21. *Sugerencia:* Considere la ecuación $A\mathbf{x} = \mathbf{0}$.

23. *Sugerencia:* Si $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial, entonces no hay variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$, y cada columna de A es una columna pivote.

25. *Sugerencia*: Considere el caso $a = b = 0$. Después considere el vector $\begin{bmatrix} -b \\ a \end{bmatrix}$, y use el hecho de que $ad - bc = 0$.

27. *Sugerencia:* Para el inciso (a), intercambie A y B en el cuadro que sigue al ejemplo 6 de la sección 2.1, y después reemplace B por la matriz identidad. Para los incisos (b) y (c), comience por escribir

$$A = \begin{bmatrix} \text{fila}_1(A) \\ \text{fila}_2(A) \\ \text{fila}_3(A) \end{bmatrix}$$

29. $\begin{bmatrix} -7 & 2 \\ 4 & -1 \end{bmatrix}$ **31.** $\begin{bmatrix} 8 & 3 & 1 \\ 10 & 4 & 1 \\ 7/2 & 3/2 & 1/2 \end{bmatrix}$

33. $A^{-1} = B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & & 0 \\ 0 & -1 & 1 & & \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$. *Sugerencia:* Para $j = 1, \ldots, n$, denote con $\mathbf{a}_j$, $\mathbf{b}_j$, y $\mathbf{e}_j$ las columnas j-ésimas de A, B e I, respectivamente. Utilice el que $\mathbf{a}_j - \mathbf{a}_{j+1} = \mathbf{e}_j$ y $\mathbf{b}_j = \mathbf{e}_j - \mathbf{e}_{j+1}$ para $j = 1, \ldots, n-1$, y $\mathbf{a}_n = \mathbf{b}_n = \mathbf{e}_n$.

35. $\begin{bmatrix} 3 \\ -6 \\ 4 \end{bmatrix}$. Encuentre esto al reducir por filas $[A \quad \mathbf{e}_3]$.

37. $C = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 0 \end{bmatrix}$

39. .27, .30 y .23 pulgadas, respectivamente.

41. **[M]** 12, 1.5, 21.5 y 12 newtons, respectivamente.

Sección 2.3, página 132

La abreviatura TMI [aquí y en la *Guía de estudio* (*Study Guide*)] denota el teorema de la matriz invertible (teorema 8).

1. Invertible, por el TMI. Ninguna de las columnas de la matriz es múltiplo de la otra columna, por lo tanto, son linealmente independientes. Además, la matriz es invertible de acuerdo con el teorema 4 de la sección 2.2 porque el determinante es diferente de cero.

3. Invertible, según el TMI. La matriz se reduce por filas a $\begin{bmatrix} 5 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ y tiene tres posiciones pivote.

5. No es invertible, según el TMI. La matriz se reduce por filas a $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & -5 \\ 0 & 0 & 0 \end{bmatrix}$ y no es equivalente por filas a I_3.

7. Invertible, de acuerdo con el TMI. La matriz se reduce por filas a $\begin{bmatrix} -1 & -3 & 0 & 1 \\ 0 & -4 & 8 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ y tiene cuatro posiciones pivote.

9. **[M]** La matriz de 4×4 tiene cuatro posiciones pivote, así que es invertible según el TMI.

11. La *Guía de estudio* (*Study Guide*) puede ayudar, pero primero trate de resolver las preguntas con base en una cuidadosa lectura del texto.

13. Una matriz cuadrada triangular superior es invertible si, y sólo si, todas las entradas de la diagonal son diferentes de cero. ¿Por qué?

Nota: Las respuestas siguientes para los ejercicios 15 a 29 mencionan al TMI. En muchos casos, parte de una respuesta aceptable (o toda ella), podría también estar basada en resultados que se usaron para establecer el TMI.

15. Si A tiene dos columnas idénticas, entonces sus columnas son linealmente dependientes. El inciso (e) del TMI muestra que A no puede ser invertible.

17. Si A es invertible, también lo es A^{-1}, según el teorema 6 de la sección 2.2. De acuerdo con (e) del TMI aplicado a A^{-1}, las columnas de A^{-1} son linealmente independientes.

19. De acuerdo con (e) del TMI, D es invertible. Así que la ecuación $D\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^7$, según (g) del TMI. ¿Qué más puede decirse?

21. La matriz G no puede ser invertible, según el teorema 5 de la sección 2.2 o el párrafo que sigue del TMI. Por lo tanto, los enunciados (g) y (h) del TMI son falsos. Las columnas de G no generan $\mathbb{R}^n$.

23. El enunciado (b) del TMI es falso para K, por lo tanto, los enunciados (e) y (h) también son falsos. Esto es, las columnas de K son linealmente *dependientes* y las columnas *no* generan $\mathbb{R}^n$.

25. *Sugerencia:* Use primero el TMI.

27. Sea W el inverso de AB. Entonces $ABW = I$ y $A(BW) = I$. Desafortunadamente, esta ecuación por sí misma no demuestra que A es invertible. ¿Por qué no? Termine la demostración antes de consultar la *Guía de estudio* (*Study Guide*).

29. Como la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no es uno a uno, el enunciado (f) del TMI es falso. Entonces (i) también es falso y la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no mapea $\mathbb{R}^n$ sobre $\mathbb{R}^n$. Además, A no es invertible, lo cual implica que la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no es invertible, según el teorema 9.

31. *Sugerencia:* Si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$, entonces A tiene un pivote en cada fila. (Teorema 4 de la sección 1.4.) ¿Podría haber variables libres en una ecuación $A\mathbf{x} = \mathbf{b}$?

33. *Sugerencia:* Primero muestre que la matriz estándar de T es invertible. Después use uno o más teoremas para mostrar que $T^{-1}(\mathbf{x}) = B\mathbf{x}$, donde $B = \begin{bmatrix} 7 & 9 \\ 4 & 5 \end{bmatrix}$.

35. *Sugerencia:* Para mostrar que T es inyectiva, suponga que $T(\mathbf{u}) = T(\mathbf{v})$ para ciertos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$. Deduzca que $\mathbf{u} = \mathbf{v}$. Para mostrar que T es suprayectiva, suponga que $\mathbf{y}$ representa un vector arbitrario en $\mathbb{R}^n$ y use el inverso S para producir un $\mathbf{x}$ tal que $T(\mathbf{x}) = \mathbf{y}$. Se puede obtener una segunda demostración usando el teorema 9 junto con un teorema de la sección 1.9.

37. *Sugerencia:* Considere las matrices estándar de T y U.

39. Dado algún $\mathbf{v}$ en $\mathbb{R}^n$, puede escribirse $\mathbf{v} = T(\mathbf{x})$ para algún $\mathbf{x}$, porque T es una función suprayectiva. Entonces las propiedades supuestas de S y U muestran que $S(\mathbf{v}) = S(T(\mathbf{x})) = \mathbf{x}$ y $U(\mathbf{v}) = U(T(\mathbf{x})) = \mathbf{x}$. Por lo tanto, $S(\mathbf{v})$ y $U(\mathbf{v})$ son iguales para todo $\mathbf{v}$. Esto es, S y U son la misma función de $\mathbb{R}^n$ en $\mathbb{R}^n$.

41. **[M]** **a.** La solución exacta de (3) es $x_1 = 3.94$ y $x_2 = .49$. La solución exacta de (4) es $x_1 = 2.90$ y $x_2 = 2.00$.

b. Cuando se utiliza la solución de (4) como aproximación a la solución de (3), el error de usar el valor 2.90 para x_1 es de aproximadamente un 26%, y el error de usar 2.0 para x_2 es de aproximadamente un 308 por ciento.

c. El número de condición de la matriz de coeficientes es 3363. El porcentaje de cambio en la solución de (3) a (4) es aproximadamente de 7700 veces el porcentaje de cambio en el miembro derecho de la ecuación. Éste es el mismo orden de magnitud que el número de condición. El número de condición proporciona una medida burda de qué tan sensible es la solución de $A\mathbf{x} = \mathbf{b}$ a cambios en $\mathbf{b}$. Se proporciona mayor información acerca del número de condición al final del capítulo 6 y en el capítulo 7.

43. **[M]** cond$(A) \approx 69{,}000$, lo cual está entre 10^4 y 10^5. Así que pueden perderse cerca de 4 o 5 dígitos de precisión. Unos cuantos experimentos con MATLAB deben verificar que $\mathbf{x}$ y $\mathbf{x}_1$ concuerdan en 11 o 12 dígitos.

45. **[M]** Algunas versiones de MATLAB ofrecen una advertencia cuando se pide invertir una matriz de Hilbert de orden 12 o mayor usando aritmética de punto flotante. El producto AA^{-1} tendrá varias entradas fuera de la diagonal que disten mucho de ser cero. Si no es así, pruebe con una matriz más grande.

Sección 2.4, página 139

1. $\begin{bmatrix} A & B \\ EA + C & EB + D \end{bmatrix}$ **3.** $\begin{bmatrix} Y & Z \\ W & X \end{bmatrix}$

5. $Y = B^{-1}$ (explique por qué), $X = -B^{-1}A$, $Z = C$

7. $X = A^{-1}$ (¿por qué?), $Y = -BA^{-1}$, $Z = 0$ (¿por qué?)

9. $X = -A_{21}A_{11}^{-1}$, $Y = -A_{31}A_{11}^{-1}$, $B_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12}$

11. Puede verificar sus respuestas en la *Guía de estudio* (*Study Guide*).

13. *Sugerencia*: Suponga que A es invertible, y sea $A^{-1} = \begin{bmatrix} D & E \\ F & G \end{bmatrix}$. Muestre que $BD = I$ y $CG = I$. Esto implica que B y C son invertibles. (Explique por qué.) De manera recíproca, suponga que B y C son invertibles. Para demostrar que A es invertible, formule una conjetura acerca de A^{-1} y compruebe que funciona.

15. $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} =$
$\begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix}$
con $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$.

17. $G_{k+1} = [\,X_k \quad \mathbf{x}_{k+1}\,]\begin{bmatrix} X_k^T \\ \mathbf{x}_{k+1}^T \end{bmatrix} = X_kX_k^T + \mathbf{x}_{k+1}\mathbf{x}_{k+1}^T$
$= G_k + \mathbf{x}_{k+1}\mathbf{x}_{k+1}^T$

Sólo se necesita calcular el producto matricial exterior $\mathbf{x}_{k+1}\mathbf{x}_{k+1}^T$ (y después sumarlo a G_k).

19. $W(s) = I_m - C(A - sI_n)^{-1}B$. Éste es el complemento de Schur de $A - sI_n$ en la matriz del sistema.

21. **a.** $A^2 = \begin{bmatrix} 1 & 0 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 3 & -1 \end{bmatrix}$

$= \begin{bmatrix} 1+0 & 0+0 \\ 3-3 & 0+(-1)^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

b. $M^2 = \begin{bmatrix} A & 0 \\ I & -A \end{bmatrix}\begin{bmatrix} A & 0 \\ I & -A \end{bmatrix}$

$= \begin{bmatrix} A^2+0 & 0+0 \\ A-A & 0+(-A)^2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

23. Si A_1 y B_1 son de $(k+1) \times (k+1)$ y triangulares inferiores, entonces puede escribirse $A_1 = \begin{bmatrix} a & \mathbf{0}^T \\ \mathbf{v} & A \end{bmatrix}$ y $B_1 = \begin{bmatrix} b & \mathbf{0}^T \\ \mathbf{w} & B \end{bmatrix}$, donde A y B son de $k \times k$ y triangulares inferiores, $\mathbf{v}$ y $\mathbf{w}$ están en $\mathbb{R}^k$, y a y b son escalares apropiados. Suponga que el producto de cualesquiera matrices triangulares inferiores de $k \times k$ es triangular inferior, y calcule el producto A_1B_1. ¿Qué puede concluirse?

25. Use el ejercicio 13 para encontrar el inverso de una matriz de la forma $B = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}$, donde B_{11} es de $p \times p$, B_{22} es de $q \times q$, y B es invertible. Divida la matriz A y aplique su resultado dos veces para encontrar que

$$A^{-1} = \begin{bmatrix} -5 & 2 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 3 & -4 \\ 0 & 0 & 0 & -5/2 & 7/2 \end{bmatrix}$$

27. **a, b.** **[M]** Los comandos a utilizar en estos ejercicios dependerán del programa de matrices.

c. El álgebra necesaria proviene de la ecuación matricial en bloques

$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}$

donde $\mathbf{x}_1$ y $\mathbf{b}_1$ están en $\mathbb{R}^{20}$ y $\mathbf{x}_2$ y $\mathbf{b}_2$ están en $\mathbb{R}^{30}$. Entonces $A_{11}\mathbf{x}_1 = \mathbf{b}_1$, que puede resolverse para producir $\mathbf{x}_1$. De la ecuación $A_{21}\mathbf{x}_1 + A_{22}\mathbf{x}_2 = \mathbf{b}_2$ se obtiene $A_{22}\mathbf{x}_2 = \mathbf{b}_2 - A_{21}\mathbf{x}_1$, que puede resolverse para $\mathbf{x}_2$ mediante la reducción por filas de la matriz $[A_{22} \quad \mathbf{c}]$, donde $\mathbf{c} = \mathbf{b}_2 - A_{21}\mathbf{x}_1$.

Sección 2.5, página 149

1. $L\mathbf{y} = \mathbf{b} \Rightarrow \mathbf{y} = \begin{bmatrix} -7 \\ -2 \\ 6 \end{bmatrix}$, $U\mathbf{x} = \mathbf{y} \Rightarrow \mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \end{bmatrix}$

3. $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -1 \\ 3 \\ 3 \end{bmatrix}$

5. $\mathbf{y} = \begin{bmatrix} 1 \\ 5 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -2 \\ -1 \\ 2 \\ -3 \end{bmatrix}$

7. $LU = \begin{bmatrix} 1 & 0 \\ -3/2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 0 & 7/2 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & 2/3 & 1 \end{bmatrix}\begin{bmatrix} 3 & -1 & 2 \\ 0 & -3 & 12 \\ 0 & 0 & -8 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1/3 & 1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -6 & 3 \\ 0 & 5 & -4 \\ 0 & 0 & 5 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 4 & 5 & 1 & 0 \\ -2 & -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 & -5 & -3 \\ 0 & -2 & 3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1/2 & -2 & 1 \end{bmatrix}\begin{bmatrix} 2 & -4 & 4 & -2 \\ 0 & 3 & -5 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$

17. $U^{-1} = \begin{bmatrix} 1/4 & 3/8 & 1/4 \\ 0 & -1/2 & 1/2 \\ 0 & 0 & 1/2 \end{bmatrix}$,

$L^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$,

$A^{-1} = \begin{bmatrix} 1/8 & 3/8 & 1/4 \\ -3/2 & -1/2 & 1/2 \\ -1 & 0 & 1/2 \end{bmatrix}$

19. *Sugerencia:* Piense en reducir por filas $[A \quad I]$.

21. *Sugerencia:* Represente las operaciones por fila mediante una sucesión de matrices elementales.

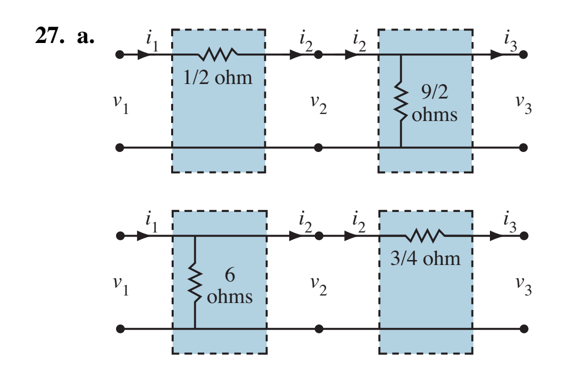
23. **a.** Denote las filas de D como transpuestas de vectores columna. Entonces la multiplicación de matrices partidas produce

$$A = CD = [\,\mathbf{c}_1 \quad \cdots \quad \mathbf{c}_4\,]\begin{bmatrix}\mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_4^T\end{bmatrix}$$
$$= \mathbf{c}_1\mathbf{v}_1^T + \cdots + \mathbf{c}_4\mathbf{v}_4^T$$

b. A tiene 40,000 entradas. Como C tiene 1600 entradas y D 400 entradas, juntas ocupan solamente el 5% de la memoria que se necesita para almacenar A.

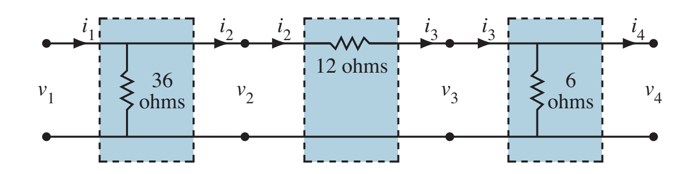
25. Explique por qué U, D y V^T son invertibles. Después use un teorema relativo al inverso de un producto de matrices invertibles.

27. a.



29. a. $\begin{bmatrix} 1 + R_2/R_1 & -R_2 \\ -1/R_1 - R_2/(R_1R_3) - 1)R_3 & 1 + R_2/R_3 \end{bmatrix}$

b. $A = \begin{bmatrix} 1 & 0 \\ -1/6 & 1 \end{bmatrix}\begin{bmatrix} 1 & -12 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1/36 & 1 \end{bmatrix}$



31. [M]

a. $$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.25 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.25 & -.0667 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -.2667 & -.2857 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -.2679 & -.0833 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -.2917 & -.2921 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -.2697 & -.0861 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -.2948 & -.2931 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 4 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.75 & -.25 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3.7333 & -1.0667 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.4286 & -.2857 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.7083 & -1.0833 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3.3919 & -.2921 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3.7052 & -1.0861 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3.3868 \end{bmatrix}$$

b. $\mathbf{x} = (3.9569, 6.5885, 4.2392, 7.3971, 5.6029, 8.7608, 9.4115, 12.0431)$

c. $$A^{-1} = \begin{bmatrix} .2953 & .0866 & .0945 & .0509 & .0318 & .0227 & .0100 & .0082 \\ .0866 & .2953 & .0509 & .0945 & .0227 & .0318 & .0082 & .0100 \\ .0945 & .0509 & .3271 & .1093 & .1045 & .0591 & .0318 & .0227 \\ .0509 & .0945 & .1093 & .3271 & .0591 & .1045 & .0227 & .0318 \\ .0318 & .0227 & .1045 & .0591 & .3271 & .1093 & .0945 & .0509 \\ .0227 & .0318 & .0591 & .1045 & .1093 & .3271 & .0509 & .0945 \\ .0100 & .0082 & .0318 & .0227 & .0945 & .0509 & .2953 & .0866 \\ .0082 & .0100 & .0227 & .0318 & .0509 & .0945 & .0866 & .2953 \end{bmatrix}$$

Obtenga A^{-1} directamente y después calcule $A^{-1} - U^{-1}L^{-1}$ para comparar los dos métodos para invertir una matriz.

Sección 2.6, página 156

1. $C = \begin{bmatrix} .10 & .60 & .60 \\ .30 & .20 & 0 \\ .30 & .10 & .10 \end{bmatrix}$, $\left\{\begin{matrix}\text{demanda} \\ \text{intermedia}\end{matrix}\right\} = \begin{bmatrix} 60 \\ 20 \\ 10 \end{bmatrix}$

3. $\mathbf{x} = \begin{bmatrix} 40 \\ 15 \\ 15 \end{bmatrix}$ **5.** $\mathbf{x} = \begin{bmatrix} 110 \\ 120 \end{bmatrix}$

7. a. $\begin{bmatrix} 1.6 \\ 1.2 \end{bmatrix}$ **b.** $\begin{bmatrix} 111.6 \\ 121.2 \end{bmatrix}$

9. $\mathbf{x} = \begin{bmatrix} 82.8 \\ 131.0 \\ 110.3 \end{bmatrix}$

11. *Sugerencia:* Use propiedades de las transpuestas para obtener $\mathbf{p}^T = \mathbf{p}^T C + \mathbf{v}^T$, de modo que $\mathbf{p}^T\mathbf{x} = (\mathbf{p}^T C + \mathbf{v}^T)\mathbf{x} = \boldsymbol{p}^T C\mathbf{x} + \mathbf{v}^T\mathbf{x}$. Ahora calcule $\boldsymbol{p}^T\mathbf{x}$ a partir de la ecuación de producción.

13. [M] $\mathbf{x} = (99576, 97703, 51231, 131570, 49488, 329554, 13835)$. Las entradas de $\mathbf{x}$ sugieren mayor precisión en la respuesta que la garantizada por las entradas en $\mathbf{d}$, las cuales aparentan ser precisas sólo hasta el millar más cercano. Entonces una respuesta más realista para $\mathbf{x}$ podría ser

$$\mathbf{x} = 1000 \times (100, 98, 51, 132, 49, 330, 14).$$

15. [M] $\mathbf{x}^{(12)}$ es el primer vector cuyas entradas son precisas hasta el millar más cercano. El cálculo de $\mathbf{x}^{(12)}$ requiere unos 1260 flops, en tanto que la reducción por filas de $[(I - C) \quad \mathbf{d}]$ requiere sólo unos 550 flops. Si C es mayor que 20×20, se necesitan menos flops para calcular $\mathbf{x}^{(12)}$ por iteración que para calcular el vector de equilibrio $\mathbf{x}$ por reducción de filas. A medida que aumenta el tamaño de C, la ventaja del método iterativo aumenta.

Asimismo, como C se vuelve más dispersa en modelos más grandes de la economía, se necesitan menos iteraciones para obtener una precisión razonable.

Sección 2.7, página 165

1. $\begin{bmatrix} 1 & .25 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ **3.** $\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & \sqrt{2} \\ \sqrt{2}/2 & \sqrt{2}/2 & 2\sqrt{2} \\ 0 & 0 & 1 \end{bmatrix}$

5. $\begin{bmatrix} \sqrt{3}/2 & 1/2 & 0 \\ 1/2 & -\sqrt{3}/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 1/2 & -\sqrt{3}/2 & 3+4\sqrt{3} \\ \sqrt{3}/2 & 1/2 & 4-3\sqrt{3} \\ 0 & 0 & 1 \end{bmatrix}$
Vea el problema de práctica.

9. $A(BD)$ requiere 1600 multiplicaciones. $(AB)D$ requiere 808 multiplicaciones. El primer método usa casi el doble de multiplicaciones. Si D tuviera 20,000 columnas, los dos conteos serían de 160,000 y 80,008, respectivamente.

11. Use el hecho de que
$$\sec\varphi - \tan\varphi \operatorname{sen}\varphi = \frac{1}{\cos\varphi} - \frac{\operatorname{sen}^2\varphi}{\cos\varphi} = \cos\varphi$$

13. $\begin{bmatrix} A & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} I & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix}\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}$. Primero aplique la transformación lineal A, y después traslade mediante $\mathbf{p}$.

15. $(12, -6, 3)$ **17.** $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 & 0 \\ 0 & \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

19. El triángulo con vértices en (7, 2, 0), (7.5, 5, 0), (5, 5, 0)

21. [M] $\begin{bmatrix} 2.2586 & -1.0395 & -.3473 \\ -1.3495 & 2.3441 & .0696 \\ .0910 & -.3046 & 1.2777 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix}$

Sección 2.8, página 173

1. El conjunto es cerrado bajo las sumas, pero no bajo la multiplicación por escalares negativos. (Esboce un ejemplo.)

3. El conjunto no es cerrado bajo las sumas ni multiplicación por escalares. El subconjunto consistente en los puntos sobre la línea $x_2 = x_1$ es un subespacio, por lo tanto, cualquier "contraejemplo" debe usar al menos un punto que no esté sobre esta línea.

5. No. El sistema correspondiente a $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{w}]$ es inconsistente.

7. **a.** Los tres vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$
b. Un número infinito de vectores
c. Sí, porque $A\mathbf{x} = \mathbf{p}$ tiene una solución.

9. No, porque $A\mathbf{p} \neq \mathbf{0}$.

11. $p = 4$ y $q = 3$. Nul A es un subespacio de $\mathbb{R}^4$ porque las soluciones de $A\mathbf{x} = \mathbf{0}$ deben tener cuatro entradas, para coincidir con las columnas de A. Col A es un subespacio de $\mathbb{R}^3$ porque cada vector columna tiene tres entradas.

13. Para Nul A, elija $(1, -2, 1, 0)$ o $(-1, 4, 0, 1)$, por ejemplo. Para Col A, seleccione cualquier columna de A.

15. Sí. Sea A la matriz cuyas columnas son los vectores dados. Entonces A es invertible porque su determinante es diferente de cero y, por lo tanto, sus columnas forman una base para $\mathbb{R}^2$, de acuerdo con el TMI (o según el ejemplo 5). (Podrían darse otras razones para la invertibilidad de A.)

17. Sí. Sea A la matriz cuyas columnas son los vectores dados. La reducción por filas de A muestra tres pivotes, así que A es invertible. Según el TMI, las columnas de A forman una base para $\mathbb{R}^3$.

19. No. Sea A la matriz de 3×2 cuyas columnas son los vectores dados. Las columnas de A no pueden generar $\mathbb{R}^3$ porque A no puede tener una posición pivote en cada fila. Así que las columnas no son una base para $\mathbb{R}^3$. (Son una base para un plano en $\mathbb{R}^3$.)

21. Lea la sección cuidadosamente y escriba sus respuestas antes de consultar la *Guía de estudio* (*Study Guide*). Esta sección contiene términos y conceptos clave que se deben aprender antes de proseguir.

23. Base para Col A: $\begin{bmatrix} 4 \\ 6 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \\ 4 \end{bmatrix}$
Base para Nul A: $\begin{bmatrix} 4 \\ -5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 6 \\ 0 \\ 1 \end{bmatrix}$

25. Base para Col A: $\begin{bmatrix} 1 \\ -1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 2 \\ 6 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ 5 \\ -5 \end{bmatrix}$
Base para Nul A: $\begin{bmatrix} 2 \\ -2.5 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ .5 \\ 0 \\ -4 \\ 1 \end{bmatrix}$

27. Construya una matriz A de 3×3 diferente de cero, y estructure $\mathbf{b}$ para que sea casi cualquier combinación lineal conveniente de las columnas de A.

29. *Sugerencia:* Es necesaria una matriz diferente de cero cuyas columnas sean linealmente dependientes.

31. Si Col $F \neq \mathbb{R}^5$, entonces las columnas de F no generan $\mathbb{R}^5$. Como F es cuadrada, el TMI muestra que F no es invertible y que la ecuación $F\mathbf{x} = \mathbf{0}$ tiene una solución no trivial. Esto es, Nul F contiene un vector diferente de cero. Otra forma de describir esto es escribir Nul $F \neq \{\mathbf{0}\}$.

33. Si Col $Q = \mathbb{R}^4$, entonces las columnas de Q generan $\mathbb{R}^4$. Como Q es cuadrada, el TMI muestra que Q es invertible y que la ecuación $Q\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^4$. Además, cada solución es única, de acuerdo con el teorema 5 de la sección 2.2.

35. Si las columnas de B son linealmente independientes, entonces la ecuación $B\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial (cero). Esto es, Nul $B = \{\mathbf{0}\}$.

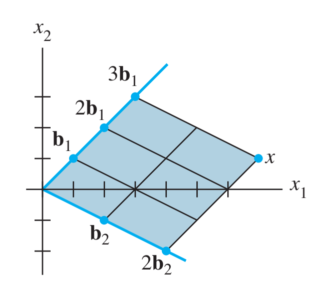
37. [**M**] Presente la forma escalonada de A, y seleccione las columnas pivote de A como una base para Col A. Para Nul A, escriba la solución de $A\mathbf{x} = \mathbf{0}$ en forma vectorial paramétrica.

Base para Col A : $\begin{bmatrix} 3 \\ -7 \\ -5 \\ 3 \end{bmatrix}, \begin{bmatrix} -5 \\ 9 \\ 7 \\ -7 \end{bmatrix}$

Base para Nul A : $\begin{bmatrix} -2.5 \\ -1.5 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4.5 \\ 2.5 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3.5 \\ -1.5 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Sección 2.9, página 180

1. $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2 = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$



3. $\begin{bmatrix} 7 \\ 5 \end{bmatrix}$ **5.** $\begin{bmatrix} 1/4 \\ -5/4 \end{bmatrix}$

7. $[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1.5 \\ .5 \end{bmatrix}$

9. Base para Col A: $\begin{bmatrix} 1 \\ -3 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -3 \\ 7 \end{bmatrix}$; dim Col $A = 3$

Base para Nul A: $\begin{bmatrix} 3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$; dim Nul $A = 1$

11. Base para Col A: $\begin{bmatrix} 1 \\ 2 \\ -3 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ -9 \\ 10 \end{bmatrix}; \begin{bmatrix} 0 \\ 4 \\ -7 \\ 11 \end{bmatrix}$; dim Col $A = 3$

Base para Nul A: $\begin{bmatrix} 9 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 0 \\ -2 \\ 1 \end{bmatrix}$; dim Nul $A = 2$

13. Las columnas 1, 3 y 4 de la matriz original forman una base para H, por lo tanto, dim $H = 3$.

15. Col $A = \mathbb{R}^3$, porque A tiene un pivote en cada fila y, por lo tanto, las columnas de A generan $\mathbb{R}^3$. Nul A *no puede* ser igual a $\mathbb{R}^2$, porque Nul A es un subespacio de $\mathbb{R}^5$. Sin embargo, es cierto que Nul A es bidimensional. Razón: la ecuación $A\mathbf{x} = \mathbf{0}$ tiene dos variables libres, porque A tiene cinco columnas y sólo tres de ellas son columnas pivote.

17. Consulte la *Guía de estudio* (*Study Guide*) después de escribir sus justificaciones.

19. El hecho de que el espacio solución de $A\mathbf{x} = \mathbf{0}$ tenga una base de tres vectores significa que dim Nul $A = 3$. Como una matriz A de 5×7 tiene siete columnas, el teorema del rango muestra que rango $A = 7 - \dim \text{Nul } A = 4$. Consulte la *Guía de estudio* (*Study Guide*) para ver una justificación que no menciona explícitamente el teorema del rango.

21. Una matriz de 7×6 tiene seis columnas. De acuerdo con el teorema del rango, dim Nul $A = 6 -$ rango A. Como el rango es cuatro, dim Nul $A = 2$. Esto es, la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$ es dos.

23. Una matriz A de 3×4 con espacio de columnas bidimensional tiene dos columnas pivote. Las dos columnas restantes corresponderán a variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$. Por lo tanto, la estructura deseada es posible. Existen seis posibles ubicaciones para las dos columnas pivote, una de las cuales es $\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Una estructura sencilla consiste en tomar dos vectores $\mathbb{R}^3$, que desde luego no sean linealmente dependientes, y colocarlos en una matriz junto con una copia de cada vector, en cualquier orden. La matriz resultante va a tener, evidentemente, un espacio de columnas bidimensional. No hay necesidad de preocuparse acerca de si Nul A tiene la dimensión correcta, puesto que eso está garantizado por el teorema del rango: dim Nul $A = 4 -$ rango A.

25. Por definición, las p columnas de A generan Col A. Si dim Col $A = p$, entonces el conjunto generador de p columnas es, de manera automática, una base para Col A, de acuerdo con el teorema de la base. En particular, las columnas son linealmente independientes.

27. **a.** *Indicación:* Las columnas de B generan W, y cada vector $\mathbf{a}_j$ está en W. El vector $\mathbf{c}_j$ está en $\mathbb{R}^p$ porque B tiene p columnas.

b. *Pista:* ¿Cuál es el tamaño de C?

c. *Pista:* ¿Cómo se relacionan B y C con A?

29. **[M]** Sus cálculos deben mostrar que la matriz $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{x}]$ corresponde a un sistema consistente. El vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$ es $(-5/3, 8/3)$.

Capítulo 2, ejercicios suplementarios, página 183

1. **a.** T **b.** F **c.** T **d.** F
e. F **f.** F **g.** T **h.** T
i. T **j.** F **k.** T **l.** F
m. F **n.** T **o.** F **p.** T

3. I

5. $A^2 = 2A - I$. Multiplique por A: $A^3 = 2A^2 - A$.
Sustituya $A^2 = 2A - I$: $A^3 = 2(2A - I) - A = 3A - 2I$.
Multiplique de nuevo por A: $A^4 = A(3A - 2I) = 3A^2 - 2A$.
Sustituya de nuevo la identidad $A^2 = 2A - I$:
$A^4 = 3(2A - I) - 2A = 4A - 3I$.

7. $\begin{bmatrix} 10 & -1 \\ 9 & 10 \\ -5 & -3 \end{bmatrix}$ **9.** $\begin{bmatrix} -3 & 13 \\ -8 & 27 \end{bmatrix}$

11. **a.**
$$\begin{aligned} p(x_i) &= c_0 + c_1x_i + \cdots + c_{n-1}x_i^{n-1} \\ &= \text{fil}_i(V)\cdot \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \end{bmatrix} = \text{fil}_i(V\mathbf{c}) = y_i \end{aligned}$$

b. Suponga que $x_1, \ldots, x_n$ son distintos, y que $V\mathbf{c} = \mathbf{0}$ para algún vector $\mathbf{c}$. Entonces las entradas de $\mathbf{c}$ son los coeficientes de un polinomio cuyo valor en los distintos puntos $x_1, \ldots, x_n$ es cero. Sin embargo, un polinomio de grado $n - 1$ diferente de cero no puede tener n ceros, así que el polinomio debe ser idénticamente cero. Es decir, las entradas de $\mathbf{c}$ deben ser todas cero. Esto demuestra que las columnas de V son linealmente independientes.

c. *Sugerencia:* Cuando $x_1, \ldots, x_n$ son distintos, existe un vector $\mathbf{c}$ tal que $V\mathbf{c} = \mathbf{y}$. ¿Por qué?

13. **a.** $P^2 = (\mathbf{u}\mathbf{u}^T)(\mathbf{u}\mathbf{u}^T) = \mathbf{u}(\mathbf{u}^T\mathbf{u})\mathbf{u}^T = \mathbf{u}(1)\mathbf{u}^T = P$
b. $P^T = (\mathbf{u}\mathbf{u}^T)^T = \mathbf{u}^{TT}\mathbf{u}^T = \mathbf{u}\mathbf{u}^T = P$

c.
$$\begin{aligned} Q^2 &= (I - 2P)(I - 2P) \\ &= I - I(2P) - 2PI + 2P(2P) \\ &= I - 4P + 4P^2 = I, \text{ por el inciso (a).} \end{aligned}$$

15. La multiplicación izquierda por una matriz elemental produce una operación elemental de fila:

$$B \sim E_1B \sim E_2E_1B \sim E_3E_2E_1B = C$$

Así que B es equivalente por filas a C. Puesto que las operaciones por fila son reversibles, C es equivalente por filas a B. (Alternativamente, muestre cómo C se transforma en B, mediante operaciones por fila, usando los inversos de las E_i.)

17. Como B es de 4×6 (con más columnas que filas), sus seis columnas son linealmente dependientes y existe un vector $\mathbf{x}$ diferente de cero tal que $B\mathbf{x} = \mathbf{0}$. Así que $A\,B\mathbf{x} = A\mathbf{0} = \mathbf{0}$, lo cual muestra que la matriz AB no es invertible, según el teorema de la matriz invertible.

19. **[M]** Hasta cuatro posiciones decimales, a medida que aumenta k,

$$A^k \to \begin{bmatrix} .2857 & .2857 & .2857 \\ .4286 & .4286 & .4286 \\ .2857 & .2857 & .2857 \end{bmatrix} \quad \text{y}$$
$$B^k \to \begin{bmatrix} .2022 & .2022 & .2022 \\ .3708 & .3708 & .3708 \\ .4270 & .4270 & .4270 \end{bmatrix}$$
o, en formato racional,
$$A^k \to \begin{bmatrix} 2/7 & 2/7 & 2/7 \\ 3/7 & 3/7 & 3/7 \\ 2/7 & 2/7 & 2/7 \end{bmatrix} \quad \text{y}$$
$$B^k \to \begin{bmatrix} 18/89 & 18/89 & 18/89 \\ 33/89 & 33/89 & 33/89 \\ 38/89 & 38/89 & 38/89 \end{bmatrix}$$

CAPÍTULO 3

Sección 3.1, página 190

1. 1 **3.** -5 **5.** -23 **7.** 4

9. 10. Inicie con la fila 3.

11. -12. Empiece con la columna 1 o la fila 4.

13. 6. Comience con la fila 2 o la columna 2.

15. 1 **17.** -5

19. $ad - bc$, $cb - da$. Intercambiar dos filas cambia el signo del determinante.

21. -2, $(18 + 12k) - (20 + 12k) = -2$. Un reemplazo de filas no cambia el valor del determinante.

23. -5, $k(4) - k(2) + k(-7) = -5k$. Escalar una fila por una constante k multiplica el determinante por k.

25. 1 **27.** k **29.** -1

31. 1. La matriz es triangular superior o inferior, con únicamente números uno en la diagonal. El determinante es 1, el producto de las entradas diagonales.

33. $\det EA = \det \begin{bmatrix} c & d \\ a & b \end{bmatrix} = cb - ad = (-1)(ad - bc)$
$= (\det E)(\det A)$

35. $\det EA = \det \begin{bmatrix} a + kc & b + kd \\ c & d \end{bmatrix}$
$= (a + kc)d - (b + kd)c$
$= ad + kcd - bc - kdc = (+1)(ad - bc)$
$= (\det E)(\det A)$

37. $5A = \begin{bmatrix} 15 & 5 \\ 20 & 10 \end{bmatrix}$; no

39. En la *Guía de estudio* (*Study Guide*) pueden encontrarse sugrencias.

41. El área del paralelogramo y el determinante de $[\mathbf{u} \quad \mathbf{v}]$ son ambos 6. Si $\mathbf{v} = \begin{bmatrix} x \\ 2 \end{bmatrix}$ para cualquier x, el área sigue siendo 6. En ningún caso cambia la base del paralelogramo, y la altura sigue siendo 2 porque la segunda coordenada de $\mathbf{v}$ es siempre 2.

43. **[M]** En general, $\det(A + B)$ no es igual a $\det A + \det B$.

45. **[M]** Podrá revisar sus conjeturas cuando llegue a la sección 3.2.

Sección 3.2, página 199

1. Intercambiar dos filas cambia el signo del determinante.

3. Una operación de reemplazo de filas no altera el determinante.

5. 3 **7.** 0 **9.** 3 **11.** 120

13. 6 **15.** 35 **17.** -7 **19.** 14

21. Invertible **23.** No invertible

25. Linealmente independiente

27. Vea la *Guía de estudio* (*Study Guide*).

29. -32

31. *Sugerencia:* Muestre que $(\det A)(\det A^{-1}) = 1$.

33. *Sugerencia:* Use el teorema 6.

35. *Sugerencia:* Use el teorema 6 y otro teorema.

37. $\det AB = \det \begin{bmatrix} 6 & 0 \\ 17 & 4 \end{bmatrix} = 24$; $(\det A)(\det B) = 3 \cdot 8 = 24$

39. **a.** -12 **b.** 500 **c.** -3 **d.** $\frac{1}{4}$ **e.** 64

41. $\det A = (a + e)d - (b + f)c = ad + ed - bc - fc$
$= (ad - bc) + (ed - fc) = \det B + \det C$

43. *Sugerencia:* Calcule $\det A$ mediante desarrollo por cofactores a lo largo de la columna 3.

45. **[M]** Consulte la *Guía de estudio* (*Study Guide*) después de haber hecho una conjetura acerca de A^TA y AA^T.

Sección 3.3, página 209

1. $\begin{bmatrix} 5/6 \\ -1/6 \end{bmatrix}$ **3.** $\begin{bmatrix} 4 \\ 5/2 \end{bmatrix}$ **5.** $\begin{bmatrix} 3/2 \\ 4 \\ -7/2 \end{bmatrix}$

7. $s \neq \pm\sqrt{3}$; $x_1 = \dfrac{5s + 4}{6(s^2 - 3)}$, $x_2 = \dfrac{-4s - 15}{4(s^2 - 3)}$

9. $s \neq 0, -1$; $x_1 = \dfrac{1}{3(s + 1)}$, $x_2 = \dfrac{4s + 3}{6s(s + 1)}$

11. $\operatorname{adj} A = \begin{bmatrix} 0 & 1 & 0 \\ -3 & -1 & -3 \\ 3 & 2 & 6 \end{bmatrix}$, $A^{-1} = \dfrac{1}{3}\begin{bmatrix} 0 & 1 & 0 \\ -3 & -1 & -3 \\ 3 & 2 & 6 \end{bmatrix}$

13. $\operatorname{adj} A = \begin{bmatrix} -1 & -1 & 5 \\ 1 & -5 & 1 \\ 1 & 7 & -5 \end{bmatrix}$, $A^{-1} = \dfrac{1}{6}\begin{bmatrix} -1 & -1 & 5 \\ 1 & -5 & 1 \\ 1 & 7 & -5 \end{bmatrix}$

15. $\operatorname{adj} A = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 6 & 0 \\ -1 & -9 & 3 \end{bmatrix}$, $A^{-1} = \dfrac{1}{6}\begin{bmatrix} 2 & 0 & 0 \\ 2 & 6 & 0 \\ -1 & -9 & 3 \end{bmatrix}$

17. Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, entonces $C_{11} = d$, $C_{12} = -c$, $C_{21} = -b$, $C_{22} = a$. La matriz adjunta es la transpuesta de cofactores:
$\operatorname{adj} A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$
Siguiendo el teorema 8, se divide entre $\det A$; esto produce la fórmua de la sección 2.2.

19. 8 **21.** 14 **23.** 22

25. Una matriz A de 3×3 no es invertible si, y sólo si, sus columnas son linealmente dependientes (de acuerdo con el teorema de la matriz invertible). Esto sucede si, y sólo si, una de las columnas está en el plano generado por las otras dos columnas, lo cual equivale a la condición de que el paralelepípedo determinado por esas columnas tenga volumen cero, lo que a su vez es equivalente a la condición de que $\det A = 0$.

27. 24 **29.** $\frac{1}{2}|\det[\mathbf{v}_1 \quad \mathbf{v}_2]|$

31. a. Vea el ejemplo 5. **b.** $4\pi abc/3$

33. [M] En MATLAB, las entradas de $B - \text{in}v(A)$ son aproximadamente 10^{-15} o más pequeñas. Consulta la *Guía de estudio* (*Study Guide*) para ver sugerencias que pueden ahorrarle golpes de tecla mientras trabaja.

35. [M] La Versión para estudiantes 4.0 de MATLAB requiere 57,771 flops para in$v(A)$, y 14,269,045 flops para la fórmula del inverso. El comando `inv(A)` requiere alrededor de sólo el 0.4% de las operaciones necesarias para la fórmula del inverso. En la *Guía de estudio* (*Study Guide*) se muestra como usar el comando `flops`.

Capítulo 3, ejercicios suplementarios, página 211

1. a. T **b.** T **c.** F **d.** F
e. F **f.** F **g.** T **h.** T
i. F **j.** F **k.** T **l.** F
m. F **n.** T **o.** F **p.** T

La solución para el ejercicio 3 se basa en el hecho de que si una matriz contiene dos filas (o dos columnas) que son múltiplos una de la otra, entonces el determinante de la matriz es cero, según el teorema 4, puesto que la matriz no puede ser invertible.

3. Efectúe dos operaciones de reemplazo de fila, y después obtenga por factorización un múltiplo común en la fila 2 y un múltiplo común en la fila 3.

$$\begin{vmatrix} 1 & a & b+c \\ 1 & b & a+c \\ 1 & c & a+b \end{vmatrix} = \begin{vmatrix} 1 & a & b+c \\ 0 & b-a & a-b \\ 0 & c-a & a-c \end{vmatrix}$$

$$= (b-a)(c-a)\begin{vmatrix} 1 & a & b+c \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{vmatrix}$$

$$= 0$$

5. -12

7. Cuando el determinante se desarrolla por cofactores de la primera fila, la ecuación tiene la forma $ax + by + c = 0$, donde al menos una variable de a y b no es cero. Ésta es la ecuación de una línea. Resulta claro que (x_1, y_1) y (x_2, y_2) están en la línea, porque cuando las coordenadas de uno de los puntos se sustituyen por x y y, dos filas de la matriz son iguales y entonces el determinante es cero.

9. $T \sim \begin{bmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{bmatrix}$. Así que, según el teorema 3,

$$\det T = (b-a)(c-a)\det\begin{bmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 1 & c+a \end{bmatrix}$$

$$= (b-a)(c-a)\det\begin{bmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 0 & c-b \end{bmatrix}$$

$$= (b-a)(c-a)(c-b)$$

11. Área = 12. Si se resta un vértice de los cuatro vértices, y si los nuevos vértices son $\mathbf{0}$, $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, entonces la figura trasladada (y por ende la figura original) será un paralelogramo si, y sólo si, $\mathbf{v}_1$, $\mathbf{v}_2$ o $\mathbf{v}_3$ es la suma de los otros dos vectores.

13. De acuerdo con la fórmula del inverso, $(\text{adj } A) \cdot \frac{1}{\det A} A = A^{-1}A = I$. Según el teorema de la matriz invertible, adj A es invertible y $(\text{adj } A)^{-1} = \frac{1}{\det A} A$.

15. a. $X = CA^{-1}$, $Y = D - CA^{-1}B$. Ahora use el ejercicio 14(c).

b. Del inciso (a) y de la propiedad multiplicativa de los determinantes,

$$\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det[A(D - CA^{-1}B)]$$
$$= \det[AD - ACA^{-1}B]$$
$$= \det[AD - CAA^{-1}B]$$
$$= \det[AD - CB]$$

donde la igualdad $AC = CA$ se utilizó en el tercer paso.

17. Primero considere el caso $n = 2$, y demuestre que el resultado es válido al calcular directamente los determinantes de B y C. Ahora suponga que la fórmula es válida para todas las matrices $(k-1) \times (k-1)$, y sean A, B y C matrices $k \times k$. Use un desarrollo por cofactores a lo largo de la primera columna y la hipótesis inductiva para encontrar det B. Utilice operaciones de reemplazo de filas sobre C para crear ceros debajo del primer pivote y producir una matriz triangular. Encuentre el determinante de esta matriz y súmelo a det B para obtener el resultado.

19. [M] Calcule:

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{vmatrix} = 1, \quad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{vmatrix} = 1,$$

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{vmatrix} = 1$$

Conjetura:

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & & 2 \\ 1 & 2 & 3 & & 3 \\ \vdots & & & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{vmatrix} = 1$$

Para confirmar la conjetura, use operaciones de reemplazo de filas para crear ceros debajo del primer pivote, después bajo el segundo pivote, y así sucesivamente. La matriz resultante es

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & & 1 \\ 0 & 0 & 1 & & 1 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}$$

que es una matriz triangular superior con determinante 1.

CAPÍTULO 4

Sección 4.1, página 223

1. a. $\mathbf{u} + \mathbf{v}$ está en V porque sus dos entradas son no negativas

b. *Ejemplo:* Si $\mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ y $c = -1$, entonces $\mathbf{u}$ está en V, pero $c\mathbf{u}$ no está en V.

3. *Ejemplo:* Si $\mathbf{u} = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$ y $c = 4$, entonces $\mathbf{u}$ está en H, pero $c\mathbf{u}$ no está en H.

5. Sí, de acuerdo con el teorema 1, puesto que el conjunto es Gen$\{t^2\}$.

7. No, el conjunto no es cerrado bajo la multiplicación por escalares que no sean enteros.

9. $H = \text{Gen }\{\mathbf{v}\}$, donde $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$. Según el teorema 1, H es un subespacio de $\mathbb{R}^3$.

11. $W = \text{Gen }\{\mathbf{u}, \mathbf{v}\}$, donde $\mathbf{u} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$. Según el teorema 1, W es un subespacio de $\mathbb{R}^3$.

13. a. Existen sólo tres vectores en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $\mathbf{w}$ no es uno de ellos.

b. Hay un número infinito de vectores en Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

c. $\mathbf{w}$ está en Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

15. No es un espacio vectorial porque el vector cero no está en W.

17. $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right\}$

19. *Sugerencia:* Use el teorema 1.

Advertencia: Aunque la *Guía de estudio* (*Study Guide*) contiene soluciones completas para todos los ejercicios impares cuyas respuestas aquí sean sólo "sugerencias", el lector *debe* tratar de encontrar la solución por sí mismo. De otro modo, no obtendrá ningún beneficio del ejercicio.

21. Sí. Las condiciones necesarias para un subespacio evidentemente se satisfacen. La matriz cero está en H, la suma de dos matrices triangulares superiores es triangular superior, y cualquier múltiplo escalar de una matriz triangular superior es, de nuevo, triangular superior.

23. Escriba sus respuestas después de leer cuidadosamente el texto.

25. 4 **27. a.** 8 **b.** 3 **c.** 5 **d.** 4

29.
$$\begin{aligned} \mathbf{u} + (-1)\mathbf{u} &= 1\mathbf{u} + (-1)\mathbf{u} && \text{Axioma 10} \\ &= [1 + (-1)]\mathbf{u} && \text{Axioma 8} \\ &= 0\mathbf{u} = \mathbf{0} && \text{Ejercicio 27} \end{aligned}$$

Del ejercicio 26, se deduce que $(-1)\mathbf{u} = -\mathbf{u}$.

31. Cualquier subespacio H que contenga $\mathbf{u}$ y $\mathbf{v}$ debe contener también todos los múltiplos escalares de $\mathbf{u}$ y $\mathbf{v}$ y, por lo tanto, todas las sumas de múltiplos escalares de $\mathbf{u}$ y $\mathbf{v}$. Entonces H debe contener a Gen$\{\mathbf{u}, \mathbf{v}\}$.

33. *Sugerencia:* Para una parte de la solución, considere $\mathbf{w}_1$ y $\mathbf{w}_2$ en $H + K$, y escriba $\mathbf{w}_1$ y $\mathbf{w}_2$ en la forma $\mathbf{w}_1 = \mathbf{u}_1 + \mathbf{v}_1$ y $\mathbf{w}_2 = \mathbf{u}_2 + \mathbf{v}_2$, donde $\mathbf{u}_1$ y $\mathbf{u}_2$ están en H, y $\mathbf{v}_1$ y $\mathbf{v}_2$ están en K.

35. [M] La forma escalonada reducida de $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \quad \mathbf{w}]$ muestra que $\mathbf{w} = 7.5\mathbf{v}_1 + 3\mathbf{v}_2 + 5.5\mathbf{v}_3$.

37. [M] Las funciones son cos $4t$ y cos $6t$. Vea el ejercicio 34 de la sección 4.5.

Sección 4.2, página 234

1. $\begin{bmatrix} 3 & -5 & -3 \\ 6 & -2 & 0 \\ -8 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, así que $\mathbf{w}$ está en Nul A.

3. $\begin{bmatrix} 7 \\ -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -6 \\ 2 \\ 0 \\ 1 \end{bmatrix}$ **5.** $\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 9 \\ 1 \\ 0 \end{bmatrix}$

7. W no es un subespacio de $\mathbb{R}^3$ porque el vector (0, 0, 0) no está en W.

9. W es un subespacio de $\mathbb{R}^4$, porque W es el conjunto de soluciones del sistema

$$\begin{aligned} a - 2b - 4c &= 0 \\ 2a - c - 3d &= 0 \end{aligned}$$

11. W no es un subespacio porque $\mathbf{0}$ no está en W. *Justificación:* Si un elemento típico $(b - 2d, 5 + d, b + 3d, d)$ fuese cero, entonces $5 + d = 0$ y $d = 0$, lo cual es imposible.

13. $W = \text{Col}\, A$ para $A = \begin{bmatrix} 1 & -6 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$, así que W es un espacio vectorial según el teorema 3.

15. $\begin{bmatrix} 0 & 2 & 3 \\ 1 & 1 & -2 \\ 4 & 1 & 0 \\ 3 & -1 & -1 \end{bmatrix}$

17. **a.** 2 **b.** 4 **19.** **a.** 5 **b.** 2

21. $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ en Nul A, $\begin{bmatrix} 2 \\ -1 \\ -4 \\ 3 \end{bmatrix}$ en Col A. Existen otras respuestas posibles.

23. $\mathbf{w}$ está en Nul A y en Col A.

25. Consulte la *Guía de estudio* (*Study Guide*), en esta etapa el lector ya debe saber cómo utilizarla.

27. Sean $\mathbf{x} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$ y $A = \begin{bmatrix} 1 & -3 & -3 \\ -2 & 4 & 2 \\ -1 & 5 & 7 \end{bmatrix}$. Entonces $\mathbf{x}$ está en Nul A. Puesto que Nul A es un subespacio de $\mathbb{R}^3$, $10\mathbf{x}$ está en Nul A.

29. **a.** $A\mathbf{0} = \mathbf{0}$, así que el vector cero está en Col A.

b. Por una propiedad de la multiplicación matricial, $A\mathbf{x} + A\mathbf{w} = A(\mathbf{x} + \mathbf{w})$, lo cual muestra que $A\mathbf{x} + A\mathbf{w}$ es una combinación lineal de las columnas de A y, por lo tanto, está en Col A.

c. $c(A\mathbf{x}) = A(c\mathbf{x})$, lo cual muestra que $c(A\mathbf{x})$ está en Col A para todo escalar c.

31. **a.** Para polinomios arbitrarios $\mathbf{p}$, $\mathbf{q}$ en $\mathbb{P}_2$ y cualquier escalar c,

$$\begin{aligned} T(\mathbf{p} + \mathbf{q}) &= \begin{bmatrix} (\mathbf{p} + \mathbf{q})(0) \\ (\mathbf{p} + \mathbf{q})(1) \end{bmatrix} = \begin{bmatrix} \mathbf{p}(0) + \mathbf{q}(0) \\ \mathbf{p}(1) + \mathbf{q}(1) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix} + \begin{bmatrix} \mathbf{q}(0) \\ \mathbf{q}(1) \end{bmatrix} = T(\mathbf{p}) + T(\mathbf{q}) \\ T(c\mathbf{p}) &= \begin{bmatrix} c\mathbf{p}(0) \\ c\mathbf{p}(1) \end{bmatrix} = c \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix} = cT(\mathbf{p}) \end{aligned}$$

Así que T es una transformación lineal de $\mathbb{P}_2$ en $\mathbb{P}_2$.

b. Cualquier polinomio cuadrático que se anula en 0 y 1 debe ser un múltiplo de $\mathbf{p}(t) = t(t - 1)$. El rango de T es $\mathbb{R}^2$.

33. **a.** Para A, B en $M_{2\times2}$ y cualquier escalar c,

$$\begin{aligned} T(A + B) &= (A + B) + (A + B)^T \\ &= A + B + A^T + B^T \quad \text{Propiedad de la transpuesta} \\ &= (A + A^T) + (B + B^T) = T(A) + T(B) \\ T(cA) &= (cA) + (cA)^T = cA + cA^T \\ &= c(A + A^T) = cT(A) \end{aligned}$$

Así que T es una transformación lineal de $M_{2\times2}$ en $M_{2\times2}$.

b. Si B es cualquier elemento en $M_{2\times2}$ con la propiedad de que $B^T = B$, y si $A = \frac{1}{2}B$, entonces

$$T(A) = \tfrac{1}{2}B + \left(\tfrac{1}{2}B\right)^T = \tfrac{1}{2}B + \tfrac{1}{2}B = B$$

c. El inciso (b) mostró que el rango de T contiene toda B tal que $B^T = B$. Así, basta probar que toda B en el rango de T tiene esta propiedad. Si $B = T(A)$, entonces, de acuerdo con las propiedades de las transpuestas,

$$B^T = (A + A^T)^T = A^T + A^{TT} = A^T + A = B$$

d. El núcleo de T es $\left\{ \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} : b \text{ real} \right\}$.

35. *Sugerencia:* Repase las tres condiciones de un subespacio. Los elementos típicos de $T(U)$ tienen la forma $T(\mathbf{u}_1)$ y $T(\mathbf{u}_2)$, donde $\mathbf{u}_1$ y $\mathbf{u}_2$ están en U.

37. **[M]** $\mathbf{w}$ está en Col A, pero no en Nul A. (Explique por qué.)

39. **[M]** La forma escalonada reducida de A es

$$\begin{bmatrix} 1 & 0 & 1/3 & 0 & 10/3 \\ 0 & 1 & 1/3 & 0 & -26/3 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Sección 4.3, página 243

1. Sí, la matriz de 3×3 $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ tiene 3 posiciones. De acuerdo con el teorema de la matriz invertible, A es es invertible y sus columnas forman una base para $\mathbb{R}^3$. (Vea el ejemplo 3.)

3. No, los vectores son linealmente dependientes y no generan $\mathbb{R}^3$.

5. No, el conjunto es linealmente dependiente porque el vector cero está en el conjunto. Sin embargo,

$$\begin{bmatrix} 1 & -2 & 0 & 0 \\ -3 & 9 & 0 & -3 \\ 0 & 0 & 0 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 3 & 0 & -3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

La matriz tiene pivotes en cada fila y, por ende, sus columnas generan $\mathbb{R}^3$.

7. No, los vectores son linealmente independientes porque no son múltiplos. (Dicho con mayor precisión, ningún vector es un múltiplo del otro.) Sin embargo, los vectores no generan $\mathbb{R}^3$. La matriz $\begin{bmatrix} -2 & 6 \\ 3 & -1 \\ 0 & 5 \end{bmatrix}$ puede tener, cuando mucho, dos pivotes puesto que sólo tiene dos columnas. Entonces no habrá un pivote en cada fila.

9. $\begin{bmatrix} 3 \\ 5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ 0 \\ 1 \end{bmatrix}$ **11.** $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

13. Bases para Nul A: $\begin{bmatrix} -6 \\ -5/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ -3/2 \\ 0 \\ 1 \end{bmatrix}$

Bases para Col A: $\begin{bmatrix} -2 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -6 \\ 8 \end{bmatrix}$

15. $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ **17.** **[M]** $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$

19. Las tres respuestas más sencillas son $\{\mathbf{v}_1, \mathbf{v}_2\}$, $\{\mathbf{v}_1, \mathbf{v}_3\}$ o $\{\mathbf{v}_2, \mathbf{v}_3\}$. Son posibles otras respuestas.

23. *Sugerencia:* Use el teorema de la matriz invertible.

25. No. (¿Por qué el conjunto no es una base para H?)

27. $\{\cos \omega t, \text{sen}\, \omega t\}$

29. Sea A la matriz $[\mathbf{v}_1 \ \cdots \ \mathbf{v}_k]$ de $n \times k$. Como A tiene menos columnas que filas, no puede haber una posición pivote en cada fila de A. De acuerdo con el teorema 4 de la sección 1.4, las columnas de A no generan $\mathbb{R}^n$ y, por lo tanto, no son una base para $\mathbb{R}^n$.

31. *Sugerencia:* Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es linealmente dependiente, entonces existen $c_1, \ldots, c_p$, no todas cero, tales que $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$. Use esta ecuación.

33. Ningún polinomio es múltiplo del otro, entonces $\{\mathbf{p}_1, \mathbf{p}_2\}$ es un conjunto linealmente independiente en $\mathbb{P}_3$.

35. Sea $\{\mathbf{v}_1, \mathbf{v}_3\}$ cualquier conjunto linealmente independiente en el espacio vectorial V, y sean $\mathbf{v}_2$ y $\mathbf{v}_4$ combinaciones lineales de $\mathbf{v}_1$ y $\mathbf{v}_3$. Entonces $\{\mathbf{v}_1, \mathbf{v}_3\}$ es una base para Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

37. **[M]** Se puede ser ingenioso y encontrar valores especiales de t que produzcan varios ceros en (5), y crear así un sistema de ecuaciones para resolver a mano con facilidad. O bien se podrían usar valores de t, como $t = 0, .1, .2, \ldots$, para crear un sistema de ecuaciones que pueda resolverse con un programa de matrices.

Sección 4.4, página 253

1. $\begin{bmatrix} 3 \\ -7 \end{bmatrix}$ **3.** $\begin{bmatrix} -1 \\ -5 \\ 9 \end{bmatrix}$ **5.** $\begin{bmatrix} 8 \\ -5 \end{bmatrix}$ **7.** $\begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}$

9. $\begin{bmatrix} 2 & 1 \\ -9 & 8 \end{bmatrix}$ **11.** $\begin{bmatrix} 6 \\ 4 \end{bmatrix}$ **13.** $\begin{bmatrix} 2 \\ 6 \\ -1 \end{bmatrix}$

15. La *Guía de estudio* (*Study Guide*) contiene sugerencias.

17. $\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5\mathbf{v}_1 - 2\mathbf{v}_2 = 10\mathbf{v}_1 - 3\mathbf{v}_2 + \mathbf{v}_3$ (un número infinito de respuestas)

19. *Pista:* Por hipótesis, el vector cero tiene una representación única como combinación lineal de elementos de S.

21. $\begin{bmatrix} 9 & 2 \\ 4 & 1 \end{bmatrix}$

23. *Sugerencia:* Suponga que $[\mathbf{u}]_{\mathcal{B}} = [\mathbf{w}]_{\mathcal{B}}$ para alguna $\mathbf{u}$ y $\mathbf{w}$ en V, y denote las entradas de $[\mathbf{u}]_{\mathcal{B}}$ con $c_1, \ldots, c_n$. Use la definición de $[\mathbf{u}]_{\mathcal{B}}$.

25. Un posible enfoque: primero, demuestre que si $\mathbf{u}_1, \ldots, \mathbf{u}_p$ son linealmente *dependientes*, entonces $[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_p]_{\mathcal{B}}$ son linealmente dependientes. Segundo, muestre que si $[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_p]_{\mathcal{B}}$ son linealmente dependientes, entonces $\mathbf{u}_1, \ldots, \mathbf{u}_p$ son linealmente *dependientes.* Use las dos ecuaciones que se muestran en el ejercicio. En la *Guía de estudio* (*Study Guide*) se proporciona una demostración poco diferente.

27. Linealmente independiente. (Justifique sus respuestas a los ejercicios 27 a 34.)

29. Linealmente dependiente

31. **a.** Los vectores de coordenadas $\begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 5 \\ -7 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -6 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ no generan $\mathbb{R}^3$. A causa del isomorfismo entre $\mathbb{R}^3$ y $\mathbb{P}_2$, los polinomios correspondientes no generan $\mathbb{P}_2$.

b. Los vectores de coordenadas $\begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -8 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}$ generan $\mathbb{R}^3$. Debido al isomorfismo entre $\mathbb{R}^3$ y $\mathbb{P}_2$, los polinomios correspondientes no generan $\mathbb{P}_2$.

33. **[M]** Los vectores de coordenadas $\begin{bmatrix} 3 \\ 7 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 16 \\ -6 \\ 2 \end{bmatrix}$ son un subconjunto linealmente dependiente en $\mathbb{R}^4$.

A causa del isomorfismo entre $\mathbb{R}^4$ y $\mathbb{P}_3$, los polinomios correspondientes forman un subconjunto linealmente dependiente en $\mathbb{P}_3$, entonces no pueden ser una base para $\mathbb{P}_3$.

35. **[M]** $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -5/3 \\ 8/3 \end{bmatrix}$ **37.** **[M]** $\begin{bmatrix} 1.3 \\ 0 \\ 0.8 \end{bmatrix}$

Sección 4.5, página 260

1. $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$; dim es 2

3. $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -3 \\ 0 \end{bmatrix}$; dim es 3

5. $\begin{bmatrix} 1 \\ 2 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ 0 \\ 7 \end{bmatrix}$; dim es 2

7. Sin base; la dimensión es 0 **9.** 2 **11.** 2 **13.** 2, 3

15. 2, 2 **17.** 0, 3 **19.** Vea la *Guía de estudio* (*Study Guide*)

21. *Pista:* Sólo es necesario demostrar que los cuatro primeros polinomios de Hermite son linealmente independientes. ¿Por qué?

23. $[\mathbf{p}]_{\mathcal{B}} = \left(3, 3, -2, \frac{3}{2}\right)$

25. *Sugerencia:* Suponga que S genera V, y aplique el teorema del conjunto generador. Esto conduce a una contradicción, lo cual demuestra que la hipótesis de generación es falsa.

27. *Sugerencia:* Utilice el hecho de que cada $\mathbb{P}_n$ es un subespacio de $\mathbb{P}$.

29. Justifique cada una de sus respuestas.

a. Verdadero **b.** Verdadero **c.** Verdadero

31. *Sugerencia:* Como H es un subespacio diferente de cero de un espacio de dimensión finita, H es de dimensión finita y tiene una base, digamos, $\mathbf{v}_1, \ldots, \mathbf{v}_p$. Primero demuestre que $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$ genera $T(H)$.

33. **[M]** **a.** Una base es $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{e}_2, \mathbf{e}_3\}$. De hecho, cualesquiera dos de los vectores $\mathbf{e}_2, \ldots, \mathbf{e}_5$ ampliarán $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ hasta una base de $\mathbb{R}^5$.

Sección 4.6, página 269

1. rango $A = 2$; $\dim \operatorname{Nul} A = 2$;

Bases para Col A: $\begin{bmatrix} 1 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ -6 \end{bmatrix}$

Bases para Fil A: $(1, 0, -1, 5), (0, -2, 5, -6)$

Bases para Nul A: $\begin{bmatrix} 1 \\ 5/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ -3 \\ 0 \\ 1 \end{bmatrix}$

3. rango $A = 3$; $\dim \operatorname{Nul} A = 2$;

Bases para Col A: $\begin{bmatrix} 2 \\ -2 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 6 \\ -3 \\ 9 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 5 \\ -4 \end{bmatrix}$

Fil A: $(2, -3, 6, 2, 5), (0, 0, 3, -1, 1), (0, 0, 0, 1, 3)$

Bases para Nul A: $\begin{bmatrix} 3/2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 9/2 \\ 0 \\ -4/3 \\ -3 \\ 1 \end{bmatrix}$

5. 5, 3, 3

7. Sí; no. Puesto que Col A es un subespacio de dimensión cuatro de $\mathbb{R}^4$, coincide con $\mathbb{R}^4$. El espacio nulo no puede ser $\mathbb{R}^3$, porque los vectores en Nul A tienen 7 entradas. Nul A es un subespacio de dimensión tres de $\mathbb{R}^7$, de acuerdo con el teorema del rango.

9. 2 **11.** 3

13. 5, 5. En ambos casos, el número de pivotes no puede exceder a la cantidad de columnas o de filas.

15. 2 **17.** Vea la *Guía de estudio* (*Study Guide*).

19. Sí. Trate de escribir una explicación antes de consultar la *Guía de estudio* (*Study Guide*).

21. No. Explique por qué.

23. Sí. Únicamente son necesarias seis ecuaciones lineales homogéneas.

25. No. Explique por qué.

27. Fil A y Nul A están en $\mathbb{R}^n$; Col A y Nul A^T están en $\mathbb{R}^m$. Sólo hay cuatro subespacios distintos porque Fil $A^T =$ Col A y Col $A^T =$ Fil A.

29. Recuerde que dim Col $A = m$ precisamente cuando Col $A = \mathbb{R}^m$ o, de manera equivalente, cuando la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para toda $\mathbf{b}$. Según el ejercicio 28(b), dim Col $A = m$ precisamente cuando dim Nul $A^T = 0$ o, de igual modo, cuando la ecuación $A^T\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial.

31. $\mathbf{u}\mathbf{v}^T = \begin{bmatrix} 2a & 2b & 2c \\ -3a & -3b & -3c \\ 5a & 5b & 5c \end{bmatrix}$. Todas las columnas son múltiplos de $\mathbf{u}$, así que Col $\mathbf{u}\mathbf{v}^T$ es unidimensional, a menos que $a = b = c = 0$.

33. *Sugerencia:* Sea $A = [\mathbf{u} \quad \mathbf{u}_2 \quad \mathbf{u}_3]$. Si $\mathbf{u} \neq \mathbf{0}$, entonces $\mathbf{u}$ es una base para Col A. ¿Por qué?

35. **[M]** *Sugerencia:* Vea el ejercicio 28 y las observaciones previas al ejemplo 4.

Sección 4.7, página 276

1. a. $\begin{bmatrix} 6 & 9 \\ -2 & -4 \end{bmatrix}$ **b.** $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$ **3.** (ii)

5. a. $\begin{bmatrix} 4 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & -2 \end{bmatrix}$ **b.** $\begin{bmatrix} 8 \\ 2 \\ 2 \end{bmatrix}$

7. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} -3 & 1 \\ -5 & 2 \end{bmatrix}$, $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} -2 & 1 \\ -5 & 3 \end{bmatrix}$

9. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 9 & -2 \\ -4 & 1 \end{bmatrix}$, $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 1 & 2 \\ 4 & 9 \end{bmatrix}$

13. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 1 & 3 & 0 \\ -2 & -5 & 2 \\ 1 & 4 & 3 \end{bmatrix}$, $[-1+2t]_{\mathcal{B}} = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$

11. Vea la *Guía de estudio* (*Study Guide*).

15. a. $\mathcal{B}$ es una base para V.

b. La función de coordenadas es una transformación lineal.

c. El producto de una matriz y un vector.

d. El vector de coordenadas de **v** relativo a $\mathcal{B}$.

17. a. **[M]**

$$P^{-1} = \frac{1}{32}\begin{bmatrix} 32 & 0 & 16 & 0 & 12 & 0 & 10 \\ & 32 & 0 & 24 & 0 & 20 & 0 \\ & & 16 & 0 & 16 & 0 & 15 \\ & & & 8 & 0 & 10 & 0 \\ & & & & 4 & 0 & 6 \\ & & & & & 2 & 0 \\ & & & & & & 1 \end{bmatrix}$$

b. P es la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$. Así que P^{-1} es la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$, de acuerdo con la ecuación (5), y las columnas de esta matriz son los vectores de $\mathcal{C}$-coordenadas de los vectores base en $\mathcal{B}$, según el teorema 15.

19. **[M]** *Sugerencia:* Sea $\mathcal{C}$ la base $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Entonces las columnas de P son $[\mathbf{u}_1]_{\mathcal{C}}$, $[\mathbf{u}_2]_{\mathcal{C}}$, y $[\mathbf{u}_3]_{\mathcal{C}}$. Use la definición de vectores de $\mathcal{C}$-coordenadas y álgebra de matrices para calcular $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$. El método de solución se analiza con la *Guía de estudio* (*Study Guide*). A continuación se presentan las respuestas numéricas:

a. $\mathbf{u}_1 = \begin{bmatrix} -6 \\ -5 \\ 21 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -6 \\ -9 \\ 32 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} -5 \\ 0 \\ 3 \end{bmatrix}$

b. $\mathbf{w}_1 = \begin{bmatrix} 28 \\ -9 \\ -3 \end{bmatrix}$, $\mathbf{w}_2 = \begin{bmatrix} 38 \\ -13 \\ 2 \end{bmatrix}$, $\mathbf{w}_3 = \begin{bmatrix} 21 \\ -7 \\ 3 \end{bmatrix}$

Sección 4.8, página 285

1. Si $y_k = 2^k$, entonces $y_{k+1} = 2^{k+1}$ y $y_{k+2} = 2^{k+2}$. Al sustituir estas fórmulas en el lado izquierdo de la ecuación se tiene

$$\begin{aligned} y_{k+2} + 2y_{k+1} - 8y_k &= 2^{k+2} + 2\cdot 2^{k+1} - 8\cdot 2^k \\ &= 2^k(2^2 + 2\cdot 2 - 8) \\ &= 2^k(0) = 0 \quad \text{para toda } k \end{aligned}$$

Como la ecuación en diferencias es válida para toda k, 2^k es una solución. Un cálculo similar funciona para $y_k = (-4)^k$.

3. Las señales 2^k y $(-4)^k$ son linealmente independientes porque ninguna es un múltiplo de la otra. Por ejemplo, no hay un escalar c tal que $2^k = c(-4)^k$ *para toda k*. De acuerdo con el teorema 17, el conjunto solución H de la ecuación en diferencias del ejercicio 1 es bidimensional. Según el teorema de la base presentado en la sección 4.5, las dos señales linealmente independientes 2^k y $(-4)^k$ forman una base para H.

5. Si $y_k = (-3)^k$, entonces

$$\begin{aligned} y_{k+2} + 6y_{k+1} + 9y_k &= (-3)^{k+2} + 6(-3)^{k+1} + 9(-3)^k \\ &= (-3)^k[(-3)^2 + 6(-3) + 9] \\ &= (-3)^k(0) = 0 \quad \text{para toda } k \end{aligned}$$

De modo similar, si $y_k = k(-3)^k$, entonces

$$\begin{aligned} &y_{k+2} + 6y_{k+1} + 9y_k \\ &\quad = (k+2)(-3)^{k+2} + 6(k+1)(-3)^{k+1} + 9k(-3)^k \\ &\quad = (-3)^k[(k+2)(-3)^2 + 6(k+1)(-3) + 9k] \\ &\quad = (-3)^k[9k + 18 - 18k - 18 + 9k] \\ &\quad = (-3)^k(0) \quad \text{para toda } k \end{aligned}$$

Así que tanto $(-3)^k$ como $k(-3)^k$ están en el espacio solución H de la ecuación en diferencias. Además, no hay un escalar c tal que $k(-3)^k = c(-3)^k$ *para toda k*, porque c debe elegirse independientemente de k. Asimismo, no hay un escalar c tal que $(-3)^k = ck(-3)^k$ para toda k. Entonces las dos señales son linealmente independientes. Como $\dim H = 2$, las señales forman una base para H, de acuerdo con el teorema de la base.

7. Sí **9.** Sí

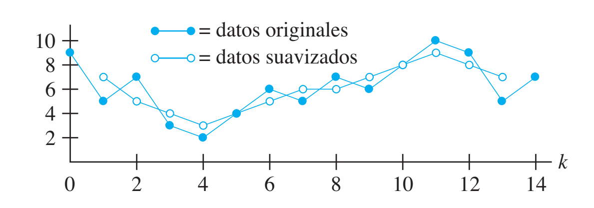
11. No, dos señales no pueden generar el espacio solución tridimensional.

13. $\left(\frac{1}{3}\right)^k, \left(\frac{2}{3}\right)^k$ **15.** $5^k, (-5)^k$

17. $Y_k = c_1(\,8)^k + c_2(\,5)^k + 10 \to 10 \quad \text{conforme } k \to \infty$

19. $y_k = c_1(-2+\sqrt{3})^k + c_2(-2-\sqrt{3})^k$

21. 7, 5, 4, 3, 4, 5, 6, 6, 7, 8, 9, 8, 7; vea la figura (en la página siguiente).



23. a. $y_{k+1} - 1.01y_k = -450,\ y_0 = 10{,}000$

25. $k^2 + c_1 \cdot (-4)^k + c_2$ **27.** $2 - 2k + c_1 \cdot 4^k + c_2 \cdot 2^{-k}$

29. $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 9 & -6 & -8 & 6 \end{bmatrix},\ \mathbf{x} = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \\ y_{k+3} \end{bmatrix}$$

31. La ecuación es valida para toda k, así que tal validez se mantiene al sustituir k por $k - 1$, lo cual transforma a la ecuación en

$y_{k+2} + 5y_{k+1} + 6y_k = 0$ para toda k

La ecuación es de orden 2.

33. Para toda k, la matriz de Casorati $C(k)$ no es invertible. En este caso, la matriz de Casorati no proporciona información acerca de la dependencia/independencia lineal del conjunto de señales. De hecho, ninguna señal es múltiplo de la otra, así que son linealmente independientes.

35. *Sugerencia:* Verifique las dos propiedades de una transformación lineal. Para $\{y_k\}$ y $\{z_k\}$ en $\mathbb{S}$, estudie $T(\{y_k\} + \{z_k\})$. Observe que si r es cualquier escalar, entonces el k-ésimo término de $r\{y_k\}$ es ry_k; así, $T(r\{y_k\})$ es la sucesión $\{w_k\}$ dada por

$w_k = ry_{k+2} + a(ry_{k+1}) + b(ry_k)$

Sección 4.9, página 296

1. a. De:

	N	M	A:
	.7	.6	Noticias
	.3	.4	Música

b. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ **c.** 33%

3. a. De:

	S	E	A:
	.95	.45	Sano
	.05	.55	Enfermo

b. 15%, 12.5%

c. .925; use $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

5. $\begin{bmatrix} .4 \\ .6 \end{bmatrix}$ **7.** $\begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix}$

9. Sí, porque P^2 tiene todas las entradas positivas.

11. a. $\begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$ **b.** 2/3

13. a. $\begin{bmatrix} .9 \\ .1 \end{bmatrix}$ **b.** .10, no

15. [M] Cerca del 13.9% de la población estadounidense.

17. a. Las entradas de una columna de P suman 1. Una columna de la matriz $P - I$ tiene las mismas entradas que P excepto que a una de las entradas se le resta 1. Por lo tanto, cada columna suma 0.

b. Según (a), la fila de abajo de $P - I$ es el negativo de la suma de las otras filas.

c. De acuerdo con (b) y el teorema del conjunto generador, la fila de abajo de $P - I$ se puede quitar y las $(n - 1)$ filas restantes aún generarán el espacio de filas. Como alternativa, use (a) y el hecho de que las operaciones por fila no cambian el espacio de filas. Sea A la matriz obtenida de $P - I$ al sumar a la última fila todas las demás filas. Según (a), el espacio de filas es generado por las primeras $(n - 1)$ filas de A.

d. De acuerdo con el teorema del rango y (c), la dimensión del espacio de columnas de $P - I$ es menor que n, así que el espacio nulo es no trivial. En lugar del teorema del rango, puede usarse el teorema de la matriz invertible, puesto que $P - I$ es una matriz cuadrada.

19. a. El producto $S\mathbf{x}$ equivale a la suma de las entradas de $\mathbf{x}$. Para un vector de probabilidad, esta suma debe ser 1.

b. $P = [\mathbf{p}_1\ \ \mathbf{p}_2\ \ \cdots\ \ \mathbf{p}_n]$, donde los $\mathbf{p}_i$ son vectores de probabilidad. De acuerdo con la multiplicación de matrices y el inciso (a).

$$SP = [\,S\mathbf{p}_1\ \ S\mathbf{p}_2\ \ \cdots\ \ S\mathbf{p}_n\,] = [\,1\ \ 1\ \ \cdots\ \ 1\,] = S$$

c. Según el inciso (b), $S(P\mathbf{x}) = (SP)\mathbf{x} = S\mathbf{x} = 1$. También, las entradas de $P\mathbf{x}$ son no negativas (porque P y $\mathbf{x}$ tienen entradas no negativas). Así, de acuerdo con (a), $P\mathbf{x}$ es un vector de probabilidad.

Capítulo 4 Ejercicios suplementarios, página 299

1. **a.** T **b.** T **c.** F **d.** F **e.** T **f.** T
g. F **h.** F **i.** T **j.** F **k.** F **l.** F
m. T **n.** F **o.** T **p.** T **q.** F **r.** T
s. T **t.** F

3. El conjunto de todos los (b_1, b_2, b_3) que satisfagan $b_1 + 2b_2 + b_3 = 0$.

5. El vector $\mathbf{p}_1$ no es cero y $\mathbf{p}_2$ no es múltiplo de $\mathbf{p}_1$, por lo que deben mantenerse ambos vectores. Como $\mathbf{p}_3 = 2\mathbf{p}_1 + 2\mathbf{p}_2$, descarte $\mathbf{p}_3$. Puesto que $\mathbf{p}_4$ tiene un término t^2, no puede ser una combinación lineal de $\mathbf{p}_1$ y $\mathbf{p}_2$, entonces conserve $\mathbf{p}_4$. Por último, $\mathbf{p}_5 = \mathbf{p}_1 + \mathbf{p}_4$, así que debe descartarse $\mathbf{p}_5$. La base resultante es $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_4\}$.

7. Tendría que saberse que el conjunto solución del sistema homogéneo es generado por dos soluciones. En este caso, el espacio nulo de la matriz de coeficientes A de 18×20 es, cuando mucho, bidimensional. Según el teorema del rango, dim Col $A \geq 20 - 2 = 18$, lo cual significa que Col $A = \mathbb{R}^{18}$, porque A tiene 18 filas y toda ecuación $A\mathbf{x} = \mathbf{b}$ es consistente.

9. Sea A la matriz estándar de $m \times n$ de la transformación T.

a. Si T es uno a uno, entonces las columnas de A son linealmente independientes (teorema 12 de la sección 1.9), así que dim Nul $A = 0$. Según el teorema del rango, dim Col A = rango $A = n$. Como el rango de T es Col A, la dimensión del rango de T es n.

b. Si T es suprayectiva, entonces las columnas de A generan $\mathbb{R}^m$ (teorema 12 de la sección 1.9), así que dim Col $A = m$. Según el teorema del rango, dim Nul $A = n -$ dim Col $A = n - m$. Como el núcleo de T es Nul A, la dimensión del núcleo de T es $n - m$.

11. Si S es un conjunto finito generador de V, entonces un subconjunto de S —por ejemplo S'— es una base de V. Puesto que S' debe generar a V, S' no puede ser un subconjunto propio de S porque S es mínimo. Por lo tanto, $S' = S$, lo cual demuestra que S' es una base para V.

12. a. *Sugerencia:* Cualquier $\mathbf{y}$ en Col AB tiene la forma $\mathbf{y} = AB\mathbf{x}$ para alguna $\mathbf{x}$.

13. Por lo visto en el ejercicio 9, rango $PA \leq$ rango A, y rango A = rango $P^{-1}PA \leq$ rango PA. Por lo tanto, rango PA = rango A.

15. La ecuación $AB = 0$ muestra que cada columna de B está en Nul A. Como Nul A es un subespacio, todas las combinaciones lineales de las columnas de B están en Nul A; por lo tanto, Col B es un subespacio de Nul A. Según el teorema 11 de la sección 4.5, dim Col $B \leq$ dim Nul A. Al aplicar el teorema del rango, se encuentra que

$$n = \text{rango } A + \text{dim Nul } A \geq \text{rango } A + \text{rango } B$$

17. a. A_1 consta de las r columnas pivote de A. Las columnas de A_1 son linealmente independientes. Por lo tanto, A_1 es una matriz de $m \times r$ con rango r.

b. De acuerdo con el teorema del rango aplicado a A_1, la dimensión de Fil A es r, por lo que A_1 tiene r filas linealmente independientes. Utilícelas para formar A_2. Entonces A_2 es de $r \times r$ con filas linealmente independientes. Según el teorema de la matriz invertible, A_2 es invertible.

19. $$[\,B \quad AB \quad A^2B\,] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -.9 & (81 \\ 1 & .5 & .25 \end{bmatrix} \sim \begin{bmatrix} 1 & -.9 & .81 \\ 0 & 1 & 0 \\ 0 & 0 & -.56 \end{bmatrix}$$

Esta matriz tiene rango menor a 3, así que el par (A, B) es controlable

21. **[M]** rango $[\,B \quad AB \quad A^2B \quad A^3B\,] = 3$. El par (A, B) no es controlable.

CAPÍTULO 5

Sección 5.1, página 308

1. Sí **3.** No **5.** Sí, $\lambda = 0$ **7.** Sí, $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

9. $\lambda = 1$: $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$; $\lambda = 5$: $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ **11.** $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$

13. $\lambda = 1$: $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$; $\lambda = 2$: $\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$; $\lambda = 3$: $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

15. $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$ **17.** $0, 2, -1$

19. 0. Justifique su respuesta.

21. Consulte la *Guía de estudio* (*Study Guide*) después de haber escrito sus respuestas.

23. *Sugerencia:* Aplique el teorema 2.

25. *Sugerencia:* Use la ecuación $A\mathbf{x} = \lambda\mathbf{x}$ para encontrar una ecuación que contenga A^{-1}.

27. *Sugerencia:* Para cualquier λ, $(A - \lambda I)^T = A^T - \lambda I$. De acuerdo con un teorema (¿cuál?), $A^T - \lambda I$ es invertible si, y sólo si, $A - \lambda I$ es invertible.

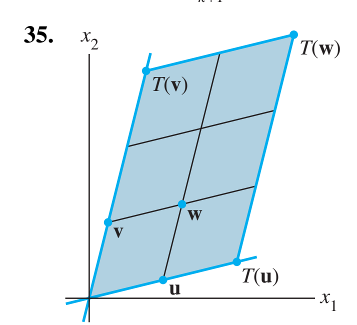
29. Sea $\mathbf{v}$ el vector en $\mathbb{R}^n$ cuyas entradas son todas números uno. Entonces $A\mathbf{v} = s\mathbf{v}$.

31. *Sugerencia:* Si A es la matriz estándar de T, busque un vector $\mathbf{v}$ diferente de cero (un punto en el plano) tal que $A\mathbf{v} = \mathbf{v}$.

33. a. $\mathbf{x}_{k+1} = c_1\lambda^{k+1}\mathbf{u} + c_2\mu^{k+1}\mathbf{v}$

b.
$$\begin{aligned} A\mathbf{x}_k &= A(c_1\lambda^k\mathbf{u} + c_2\mu^k\mathbf{v}) \\ &= c_1\lambda^k A\mathbf{u} + c_2\mu^k A\mathbf{v} \quad \text{Linealidad} \\ &= c_1\lambda^k \lambda\mathbf{u} + c_2\mu^k \mu\mathbf{v} \quad \mathbf{u} \text{ y } \mathbf{v} \text{ son vectores propios.} \\ &= \mathbf{x}_{k+1} \end{aligned}$$

35.



37. **[M]** $\lambda = 3$: $\begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$; $\lambda = 13$: $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Se pueden acelerar los cálculos con el programa `nulbasis`, que se analiza en la *Guía de estudio* (*Study Guide*).

39. **[M]** $\lambda = -2$: $\begin{bmatrix} -2 \\ 7 \\ -5 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 7 \\ -5 \\ 0 \\ 5 \end{bmatrix}$;

$\lambda = 5$: $\begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Sección 5.2, página 317

1. $\lambda^2 - 4\lambda - 45$; 9, −5 **3.** $\lambda^2 - 2\lambda - 1$; $1 \pm \sqrt{2}$

5. $\lambda^2 - 6\lambda + 9$; 3

7. $\lambda^2 - 9\lambda + 32$; no hay valores propios reales

9. $-\lambda^3 + 4\lambda^2 - 9\lambda - 6$ **11.** $-\lambda^3 + 9\lambda^2 - 26\lambda + 24$

13. $-\lambda^3 + 18\lambda^2 - 95\lambda + 150$ **15.** 4, 3, 3, 1

17. 3, 3, 1, 1, 0

19. *Sugerencia:* La ecuación dada es válida para toda λ.

21. En la *Guía de estudio* (*Study Guide*) se incluyen sugerencias.

23. *Sugerencia:* Encuentre una matriz P invertible tal que $RQ = P^{-1}AP$.

25. **a.** $\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ es un vector propio para $\lambda = .3$

b. $\mathbf{x}_0 = \mathbf{v}_1 - \frac{1}{14}\mathbf{v}_2$

c. $\mathbf{x}_1 = \mathbf{v}_1 - \frac{1}{14}(.3)\mathbf{v}_2$, $\mathbf{x}_2 = \mathbf{v}_1 - \frac{1}{14}(.3)^2\mathbf{v}_2$, y $\mathbf{x}_k = \mathbf{v}_1 - \frac{1}{14}(.3)^k\mathbf{v}_2$. Conforme $k \to \infty$, $(.3)^k \to 0$ y $\mathbf{x}_k \to \mathbf{v}_1$.

27. **a.** $A\mathbf{v}_1 = \mathbf{v}_1$, $A\mathbf{v}_2 = .5\mathbf{v}_2$, $A\mathbf{v}_3 = .2\mathbf{v}_3$. (Esto muestra también que los valores propios de A son 1, .5, y .2.)

b. $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente porque los vectores propios corresponden a valores propios distintos (teorema 2). Como hay tres vectores en el conjunto, éste es una base para $\mathbb{R}^3$. Así, existen constantes (únicas) tales que

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$

Entonces

$$\mathbf{w}^T\mathbf{x}_0 = c_1\mathbf{w}^T\mathbf{v}_1 + c_2\mathbf{w}^T\mathbf{v}_2 + c_3\mathbf{w}^T\mathbf{v}_3 \qquad (*)$$

Dado que $\mathbf{x}_0$ y $\mathbf{v}_1$ son vectores de probabilidad y como las entradas de $\mathbf{v}_2$ y $\mathbf{v}_3$ en cada caso suman 0, (*) muestra que $1 = c_1$.

c. De acuerdo con (b),

$$\mathbf{x}_0 = \mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$

Al aplicar (a),

$$\begin{aligned} \mathbf{x}_k &= A^k\mathbf{x}_0 = A^k\mathbf{v}_1 + c_2A^k\mathbf{v}_2 + c_3A^k\mathbf{v}_3 \\ &= \mathbf{v}_1 + c_2(\ 5)^k\mathbf{v}_2 + c_3(.2)^k\mathbf{v}_3 \\ &\to \mathbf{v}_1 \text{ conforme } k \to \infty \end{aligned}$$

29. **[M]** Informe acerca de sus resultados y conclusiones. Es posible evitar los cálculos tediosos si se utiliza el programa `gauss`, el cual se analiza en la *Guía de estudio* (*Study Guide*).

Sección 5.3, página 325

1. $\begin{bmatrix} 226 & -525 \\ 90 & -209 \end{bmatrix}$ **3.** $\begin{bmatrix} a^k & 0 \\ 3(a^k - b^k) & b^k \end{bmatrix}$

5. $\lambda = 5$: $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$; $\lambda = 1$: $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$

Cuando una respuesta implique diagonalización, $A = PDP^{-1}$, los factores P y D no son únicos, así que las respuestas pueden diferir de las proporcionadas aquí.

7. $P = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ **9.** No diagonalizable

11. $P = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 3 & 1 \\ 4 & 3 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

13. $P = \begin{bmatrix} -1 & 2 & 1 \\ -1 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

15. $P = \begin{bmatrix} -1 & -4 & -2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

17. No diagonalizable

19. $P = \begin{bmatrix} 1 & 3 & -1 & -1 \\ 0 & 2 & -1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

21. Vea la *Guía de estudio* (*Study Guide*).

23. Sí. (Explique por qué.)

25. No, A debe ser diagonalizable. (Explique por qué.)

27. *Sugerencia:* Escriba $A = PDP^{-1}$. Como A es invertible, 0 no es un valor propio de A, así que D tiene entradas diferentes de cero en su diagonal.

29. Una respuesta es $P_1 = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix}$, cuyas columnas son vectores propios correspondientes a los valores propios en D_1.

31. *Sugerencia:* Construya una apropiada matriz triangular de 2×2.

33. [M] $P = \begin{bmatrix} 2 & 2 & 1 & 6 \\ 1 & -1 & 1 & -3 \\ -1 & -7 & 1 & 0 \\ 2 & 2 & 0 & 4 \end{bmatrix}$,

$D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$

35. [M] $P = \begin{bmatrix} 6 & 3 & 2 & 4 & 3 \\ -1 & -1 & -1 & -3 & -1 \\ -3 & -3 & -4 & -2 & -4 \\ 3 & 0 & -1 & 5 & 0 \\ 0 & 3 & 4 & 0 & 5 \end{bmatrix}$,

$D = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Sección 5.4, página 333

1. $\begin{bmatrix} 3 & -1 & 0 \\ -5 & 6 & 4 \end{bmatrix}$

3. a. $T(\mathbf{e}_1) = -\mathbf{b}_2 + \mathbf{b}_3$, $T(\mathbf{e}_2) = -\mathbf{b}_1 - \mathbf{b}_3$, $T(\mathbf{e}_3) = \mathbf{b}_1 - \mathbf{b}_2$

b. $[T(\mathbf{e}_1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$, $[T(\mathbf{e}_2)]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}$,

$[T(\mathbf{e}_3)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

c. $\begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$

5. a. $10 - 3t + 4t^2 + t^3$

b. Para cualesquier $\mathbf{p}$, $\mathbf{q}$ en $\mathbb{P}_2$ y cualquier escalar c,

$$\begin{aligned} T[\mathbf{p}(t) + \mathbf{q}(t)] &= (t+5)[\mathbf{p}(t) + \mathbf{q}(t)] \\ &= (t+5)\mathbf{p}(t) + (t+5)\mathbf{q}(t) \\ &= T[\mathbf{p}(t)] + T[\mathbf{q}(t)] \\ T[c\cdot\mathbf{p}(t)] &= (t+5)[c\cdot\mathbf{p}(t)] = c\cdot(t+5)\mathbf{p}(t) \\ &= c\cdot T[\mathbf{p}(t)] \end{aligned}$$

c. $\begin{bmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 3 & 0 & 0 \\ 5 & -2 & 0 \\ 0 & 4 & 1 \end{bmatrix}$

9. a. $\begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$

b. *Sugerencia:* Calcule $T(\mathbf{p} + \mathbf{q})$ y $T(c \cdot \mathbf{p})$ para $\mathbf{p}$, $\mathbf{q}$ arbitrarias en $\mathbb{P}_2$ y un escalar arbitrario c.

c. $\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}$ **13.** $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

15. $\mathbf{b}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

17. a. $A\mathbf{b}_1 = 2\mathbf{b}_1$, así que $\mathbf{b}_1$ es un vector propio de A. Sin embargo, A tiene sólo un valor propio, $\lambda = 2$, y el espacio propio solamente es unidimensional, de modo que A no es diagonalizable.

b. $\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$

19. Por definición, si A es semejante a B, existe una matriz invertible P tal que $P^{-1}AP = B$. (Consulte la sección 5.2.) Entonces B es invertible porque es el producto de matrices invertibles. Para demostrar que A^{-1} es semejante a B^{-1}, use la ecuación $P^{-1}AP = B$. Vea la *Guía de estudio* (*Study Guide*).

21. *Sugerencia:* Repase el problema de práctica 2.

23. *Sugerencia:* Calcule $B(P^{-1}\mathbf{x})$.

25. *Sugerencia:* Escriba $A = PBP^{-1} = (PB)P^{-1}$, y use la propiedad de la traza.

27. Para cada j, $I(\mathbf{b}_j) = \mathbf{b}_j$. Puesto que el vector de coordenadas estándar de cualquier vector en $\mathbb{R}^n$ es propiamente el mismo vector, $[I(\mathbf{b}_j)]_{\mathcal{E}} = \mathbf{b}_j$. Por lo tanto, la matriz para I relativa a $\mathcal{B}$ y la base estándar $\mathcal{E}$ es simplemente $[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n]$. Esta matriz es precisamente la matriz de *cambio de coordenadas* $P_{\mathcal{B}}$ que se definió en la sección 4.4.

29. La $\mathcal{B}$-matriz de la transformación de identidad es I_n, porque el vector de $\mathcal{B}$-coordenadas del j-ésimo vector de base $\mathbf{b}_j$ es la j-ésima columna de I_n.

31. [M] $\begin{bmatrix} -7 & -2 & -6 \\ 0 & -4 & -6 \\ 0 & 0 & -1 \end{bmatrix}$

Sección 5.5, página 341

1. $\lambda = 2 + i, \begin{bmatrix} -1+i \\ 1 \end{bmatrix}; \quad \lambda = 2 - i, \begin{bmatrix} -1-i \\ 1 \end{bmatrix}$

3. $\lambda = 2 + 3i, \begin{bmatrix} 1-3i \\ 2 \end{bmatrix}; \quad \lambda = 2 - 3i, \begin{bmatrix} 1+3i \\ 2 \end{bmatrix}$

5. $\lambda = 2 + 2i, \begin{bmatrix} 1 \\ 2+2i \end{bmatrix}; \quad \lambda = 2 - 2i, \begin{bmatrix} 1 \\ 2-2i \end{bmatrix}$

7. $\lambda = \sqrt{3} \pm i$, $\varphi = \pi/6$ radianes, $r = 2$

9. $\lambda = -\sqrt{3}/2 \pm (1/2)i$, $\varphi = -5\pi/6$ radianes, $r = 1$

11. $\lambda = .1 \pm .1i$, $\varphi = -\pi/4$ radianes, $r = \sqrt{2}/10$

En los ejercicios 13 a 20, hay otras posibles respuestas. Cualquier P que vuelva $P^{-1}AP$ igual a la C dada o a C^T es una respuesta satisfactoria. Primero encuentre P; después calcule $P^{-1}AP$.

13. $P = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$

15. $P = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}, C = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$

17. $P = \begin{bmatrix} 2 & -1 \\ 5 & 0 \end{bmatrix}, C = \begin{bmatrix} -.6 & -.8 \\ .8 & -.6 \end{bmatrix}$

19. $P = \begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}, C = \begin{bmatrix} .96 & -.28 \\ .28 & .96 \end{bmatrix}$

21. $\mathbf{y} = \begin{bmatrix} 2 \\ -1+2i \end{bmatrix} = \dfrac{-1+2i}{5} \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$

23. (a) Propiedades de las conjugadas y el hecho de que $\overline{\mathbf{x}}^T = \overline{\mathbf{x}^T}$; (b) $\overline{A\mathbf{x}} = A\overline{\mathbf{x}}$ y A es real; (c) porque $\mathbf{x}^T A\overline{\mathbf{x}}$ es un escalar y, por lo tanto, puede verse como una matriz de 1×1; (d) propiedades de las transpuestas; (e) $A^T = A$, definición de q.

25. *Sugerencia:* Primero escriba $\mathbf{x} = \text{Re}\,\mathbf{x} + i(\text{Im}\,\mathbf{x})$.

27. [M] $P = \begin{bmatrix} 1 & -1 & -2 & 0 \\ -4 & 0 & 0 & 2 \\ 0 & 0 & -3 & -1 \\ 2 & 0 & 4 & 0 \end{bmatrix}$,

$C = \begin{bmatrix} .2 & -.5 & 0 & 0 \\ .5 & .2 & 0 & 0 \\ 0 & 0 & .3 & -.1 \\ 0 & 0 & .1 & .3 \end{bmatrix}$

Son posibles otras opciones, pero C debe ser igual a $P^{-1}AP$.

Sección 5.6, página 352

1. a. *Sugerencia:* Encuentre c_1, c_2 tales que $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Use esta representación, y el hecho de que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios de A, para calcular $\mathbf{x}_1 = \begin{bmatrix} 49/3 \\ 41/3 \end{bmatrix}$.

b. En general, $\mathbf{x}_k = 5(3)^k\mathbf{v}_1 - 4(\frac{1}{3})^k\mathbf{v}_2$ para $k \geq 0$.

3. Cuando $p = .2$, los valores propios de A son .9 y .7, y

$$\mathbf{x}_k = c_1(.9)^k \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2(.7)^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} \to \mathbf{0} \quad \text{conforme } k \to \infty$$

La mayor tasa de depredación disminuye el abasto de comida del búho, y tarde o temprano tanto la población del depredador como la de la presa perecen.

5. Si $p = .325$, los valores propios son 1.05 y .55. Puesto que $1.05 > 1$, ambas poblaciones crecerán un 5% al año. Un vector propio para 1.05 es (6, 13), así que tarde o temprano habrá 6 búhos manchados por cada 13 (mil) ardillas voladoras.

7. a. El origen es un punto silla porque A tiene un valor propio mayor que 1 y uno menor que 1 (en valor absoluto).

b. La dirección de mayor atracción está dada por el vector propio correspondiente al valor propio de $1/3$, a saber, $\mathbf{v}_2$. Todos los vectores que son múltiplos de $\mathbf{v}_2$ son atraídos al origen. La dirección de mayor repulsión está dada por el vector propio $\mathbf{v}_1$. Todos los múltiplos de $\mathbf{v}_1$ son repelidos.

c. Vea la *Guía de estudio* (*Study Guide*).

9. Punto silla; valores propios: 2, .5; dirección de mayor repulsión: la línea que pasa por (0, 0) y (−1, 1); dirección de mayor atracción: la línea que pasa por (0, 0) y (1, 4).

11. Atractor; valores propios: .9, .8: mayor atracción: la línea que pasa por (0, 0) y (5, 4).

13. Repulsor; valores propios: 1.2, 1.1; mayor repulsión: la línea que pasa por (0, 0) y (3, 4).

15. $\mathbf{x}_k = \mathbf{v}_1 + .1(.5)^k \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} + .3(.2)^k \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \to \mathbf{v}_1$ conforme $k \to \infty$

17. **a.** $A = \begin{bmatrix} 0 & 1.6 \\ .3 & .8 \end{bmatrix}$

b. La población crece porque el mayor valor propio de A es 1.2, cuya magnitud es mayor que 1. La tasa de crecimiento final es de 1.2, lo cual significa un 20% anual. El vector propio (4, 3) para $\lambda_1 = 1.2$ muestra que habrá 4 juveniles por cada 3 adultos.

c. **[M]** La proporción juveniles-adultos parece estabilizarse después de 5 o 6 años. La *Guía de estudio* (*Study Guide*) describe como construir un programa de matrices para generar una matriz de datos cuyas columnas enlisten los números de jóvenes y adultos cada año. También se analiza la gratificación de los datos.

Sección 5.7, página 361

1. $\mathbf{x}(t) = \frac{5}{2}\begin{bmatrix} -3 \\ 1 \end{bmatrix} e^{4t} - \frac{3}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{2t}$

3. $-\frac{5}{2}\begin{bmatrix} -3 \\ 1 \end{bmatrix} e^{t} + \frac{9}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-t}$. El origen es un punto silla. La dirección de mayor atracción es la línea que pasa por $(-1, 1)$ y el origen. La dirección de mayor repulsión es la línea que pasa por $(-3, 1)$ y el origen.

5. $-\frac{1}{2}\begin{bmatrix} 1 \\ 3 \end{bmatrix} e^{4t} + \frac{7}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{6t}$. El origen es un repulsor. La dirección de mayor repulsión es la línea que pasa por (1, 1) y el origen.

7. Sea $P = \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}$ y $D = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}$. Entonces $A = PDP^{-1}$.

Sustituyendo $\mathbf{x} = P\mathbf{y}$ en $\mathbf{x}' = A\mathbf{x}$, se tiene que

$$\begin{aligned} \frac{d}{dt}(P\mathbf{y}) &= A(P\mathbf{y}) \\ P\mathbf{y}' &= PDP^{-1}(P\mathbf{y}) = PD\mathbf{y} \end{aligned}$$

La multiplicación izquierda por P^{-1} da

$$\mathbf{y}' = D\mathbf{y}, \quad \text{o bien} \quad \begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

9. (solución compleja):

$$c_1 \begin{bmatrix} 1-i \\ 1 \end{bmatrix} e^{(-2+i)t} + c_2 \begin{bmatrix} 1+i \\ 1 \end{bmatrix} e^{(-2-i)t}$$

(solución real):

$$c_1 \begin{bmatrix} \cos t + \operatorname{sen} t \\ \cos t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \operatorname{sen} t - \cos t \\ \operatorname{sen} t \end{bmatrix} e^{-2t}$$

Las trayectorias forman una espiral hacia el origen.

11. (compleja): $c_1 \begin{bmatrix} -3+3i \\ 2 \end{bmatrix} e^{3it} + c_2 \begin{bmatrix} -3-3i \\ 2 \end{bmatrix} e^{-3it}$

(real):

$$c_1 \begin{bmatrix} -3\cos 3t - 3\operatorname{sen} 3t \\ 2\cos 3t \end{bmatrix} + c_2 \begin{bmatrix} -3\operatorname{sen} 3t + 3\cos 3t \\ 2\operatorname{sen} 3t \end{bmatrix}$$

Las trayectorias son elipses alrededor del origen.

13. (compleja): $c_1 \begin{bmatrix} 1+i \\ 2 \end{bmatrix} e^{(1+3i)t} + c_2 \begin{bmatrix} 1-i \\ 2 \end{bmatrix} e^{(1-3i)t}$

(real): $c_1 \begin{bmatrix} \cos 3t - \operatorname{sen} 3t \\ 2\cos 3t \end{bmatrix} e^{t} + c_2 \begin{bmatrix} \operatorname{sen} 3t + \cos 3t \\ 2\operatorname{sen} 3t \end{bmatrix} e^{t}$

Las trayectorias se alejan en espiral del origen.

15. **[M]** $\mathbf{x}(t) = c_1 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} -6 \\ 1 \\ 5 \end{bmatrix} e^{-t} + c_3 \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} e^{t}$

El origen es un punto silla. Una solución con $c_3 = 0$ es atraída al origen. Una solución con $c_1 = c_2 = 0$ es repelida.

17. **[M]** (compleja):

$$c_1 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} e^{t} + c_2 \begin{bmatrix} 23-34i \\ -9+14i \\ 3 \end{bmatrix} e^{(5+2i)t} + c_3 \begin{bmatrix} 23+34i \\ -9-14i \\ 3 \end{bmatrix} e^{(5-2i)t}$$

(real): $c_1 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} e^{t} + c_2 \begin{bmatrix} 23\cos 2t + 34\operatorname{sen} 2t \\ -9\cos 2t - 14\operatorname{sen} 2t \\ 3\cos 2t \end{bmatrix} e^{5t} +$

$$c_3 \begin{bmatrix} 23\operatorname{sen} 2t - 34\cos 2t \\ -9\operatorname{sen} 2t + 14\cos 2t \\ 3\operatorname{sen} 2t \end{bmatrix} e^{5t}$$

El origen es un repulsor. Las trayectorias se alejan en espiral del origen.

19. **[M]** $A = \begin{bmatrix} -2 & 3/4 \\ 1 & -1 \end{bmatrix}$,

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \frac{5}{2}\begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} - \frac{1}{2}\begin{bmatrix} -3 \\ 2 \end{bmatrix} e^{-2.5t}$$

21. **[M]** $A = \begin{bmatrix} -1 & -8 \\ 5 & -5 \end{bmatrix}$,

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -20\operatorname{sen} 6t \\ 15\cos 6t - 5\operatorname{sen} 6t \end{bmatrix} e^{-3t}$$

Sección 5.8, página 368

1. Vecor propio: $\mathbf{x}_4 = \begin{bmatrix} 1 \\ .3326 \end{bmatrix}$, o bien $A\mathbf{x}_4 = \begin{bmatrix} 4.9978 \\ 1.6652 \end{bmatrix}$; $\lambda \approx 4.9978$

3. Vector propio: $\mathbf{x}_4 = \begin{bmatrix} .5188 \\ 1 \end{bmatrix}$, o bien $A\mathbf{x}_4 = \begin{bmatrix} .4594 \\ .9075 \end{bmatrix}$; $\lambda \approx .9075$

5. $\mathbf{x} = \begin{bmatrix} -.7999 \\ 1 \end{bmatrix}$, $A\mathbf{x} = \begin{bmatrix} 4.0015 \\ -5.0020 \end{bmatrix}$; estimado $\lambda = -5.0020$

7. **[M]** $\mathbf{x}_k$: $\begin{bmatrix} .75 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ .9565 \end{bmatrix}, \begin{bmatrix} .9932 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ .9990 \end{bmatrix}, \begin{bmatrix} .9998 \\ 1 \end{bmatrix}$

μ_k: 11.5, 12.78, 12.96, 12.9948, 12.9990

9. **[M]** $\mu_5 = 8.4233$, $\mu_6 = 8.4246$; valor real: 8.42443 (preciso hasta 5 decimales)

11. μ_k: 5.8000, 5.9655, 5.9942, 5.9990 $(k = 1, 2, 3, 4)$; $R(\mathbf{x}_k)$: 5.9655, 5.9990, 5.99997, 5.9999993

13. Sí, pero las sucesiones podrían converger lentamente.

15. *Sugerencia:* Escriba $A\mathbf{x} - \alpha\mathbf{x} = (A - \alpha I)\mathbf{x}$, y use el hecho de que $(A - \alpha I)$ es invertible cuando α *no* es un valor propio de A.

17. **[M]** $v_0 = 3.3384$, $v_1 = 3.32119$ (preciso hasta 4 decimales con redondeo), $v_2 = 3.3212209$. Valor real: 3.3212201 (preciso hasta 7 decimales)

19. **a.** $\mu_6 = 30.2887 = \mu_7$ hasta cuatro decimales. Hasta seis decimales, el mayor valor propio es 30.288685, con vector propio (.957629, .688937, 1, .943782).

b. El método de la potencia inversa (con $\alpha = 0$) produce $\mu_1^{-1} = .010141$, $\mu_2^{-1} = .010150$. Hasta siete decimales, el menor valor propio es .0101500, con vector propio (−.603972, 1, −.251135, .148953). La razón de la convergencia rápida es que el valor propio que sigue al menor está cerca de .85.

21. **a.** Si los valores propios de A son todos de magnitud menor que 1, y si $\mathbf{x} \neq \mathbf{0}$, entonces $A^k\mathbf{x}$ es aproximadamente un vector propio para k grande.

b. Si el valor propio estrictamente dominante es 1, y si $\mathbf{x}$ tiene una componente en la dirección del vector propio correspondiente, entonces $\{A^k\mathbf{x}\}$ convergerá a un múltiplo de dicho vector propio.

c. Si los valores propios de A son todos mayores en magnitud que 1, y si $\mathbf{x}$ no es un vector propio, entonces la distancia de $A^k\mathbf{x}$ al vector propio más cercano *aumentará* conforme $k \to \infty$.

Capítulo 5 Ejercicios suplementarios, página 370

1. **a.** T **b.** F **c.** T **d.** F **e.** T
f. T **g.** F **h.** T **i.** F **j.** T
k. F **l.** F **m.** F **n.** T **o.** F
p. T **q.** F **r.** T **s.** F **t.** T
u. T **v.** T **w.** F **x.** T

3. **a.** Suponga que $A\mathbf{x} = \lambda\mathbf{x}$, con $\mathbf{x} \neq \mathbf{0}$. Entonces $(5I - A)\mathbf{x} = 5\mathbf{x} - A\mathbf{x} = 5\mathbf{x} - \lambda\mathbf{x} = (5 - \lambda)\mathbf{x}$. El valor propio es $5 - \lambda$.

b. $(5I - 3A + A^2)\mathbf{x} = 5\mathbf{x} - 3A\mathbf{x} + A(A\mathbf{x}) = 5\mathbf{x} - 3\lambda\mathbf{x} + \lambda^2\mathbf{x} = (5 - 3\lambda + \lambda^2)\mathbf{x}$. El valor propio es $5 - 3\lambda + \lambda^2$.

5. Suponga que $A\mathbf{x} = \lambda\mathbf{x}$, con $\mathbf{x} \neq \mathbf{0}$. Entonces

$$\begin{aligned} p(A)\mathbf{x} &= (c_0 I + c_1 A + c_2 A^2 + \cdots + c_n A^n)\mathbf{x} \\ &= c_0\mathbf{x} + c_1 A\mathbf{x} + c_2 A^2\mathbf{x} + \cdots + c_n A^n\mathbf{x} \\ &= c_0\mathbf{x} + c_1\lambda\mathbf{x} + c_2\lambda^2\mathbf{x} + \cdots + c_n\lambda^n\mathbf{x} = p(\lambda)\mathbf{x} \end{aligned}$$

Así que $p(\lambda)$ es un valor propio de $p(A)$.

7. Si $A = PDP^{-1}$, entonces $p(A) = Pp(D)P^{-1}$, como se muestra en el ejercicio 6. Si la entrada (j, j) en D es λ, entonces la entrada (j, j) de D^k es λ^k, y entonces la entrada (j, j) de $p(D)$ es $p(\lambda)$. Si p es el polinomio característico de A, entonces $p(\lambda) = 0$ para toda entrada diagonal en D, porque estas entradas en D son los valores propios de A. Por lo tanto, $p(D)$ es la matriz cero, y $p(A) = P \cdot 0 \cdot P^{-1} = 0$.

9. Si $I - A$ fuera no invertible, entonces la ecuación $(I - A)\mathbf{x} = \mathbf{0}$ tendría una solución no trivial $\mathbf{x}$. Entonces $\mathbf{x} - A\mathbf{x} = \mathbf{0}$ y $A\mathbf{x} = 1 \cdot \mathbf{x}$, lo cual demuestra que A tendría a 1 como valor propio. Esto no puede suceder si todos los valores propios son de magnitud menor que 1. Así que $I - A$ debe ser invertible.

11. **a.** Tome $\mathbf{x}$ en H. Entonces $\mathbf{x} = c\mathbf{u}$ para algún escalar c. Entonces $A\mathbf{x} = A(c\mathbf{u}) = c(A\mathbf{u}) = c(\lambda\mathbf{u}) = (c\lambda)\mathbf{u}$, lo cual muestra que $A\mathbf{x}$ está en H.

b. Sea $\mathbf{x}$ un vector diferente de cero en K. Puesto que K es unidimensional, K debe ser el conjunto de todos los múltiplos escalares de $\mathbf{x}$. Si K es invariante bajo A, entonces $A\mathbf{x}$ está en K y, por lo tanto, $A\mathbf{x}$ es múltiplo de $\mathbf{x}$. Por consiguiente, $\mathbf{x}$ es un vector propio de A.

13. 1, 3, 7

15. Reemplace a por $a - \lambda$ en la fórmula del determinante presentada en el ejercicio 16 del capítulo 3 en los ejercicios suplementarios:

$$\det(A - \lambda I) = (a - b - \lambda)^{n-1}[a - \lambda +)n - 1)b]$$

Este determinante es cero sólo si $a - b - \lambda = 0$ o $a - \lambda + (n - 1)b = 0$. Por lo tanto, λ es un valor propio de A si, y sólo si, $\lambda = a - b$ o $\lambda = a + (n - 1)$. A partir de la fórmula para $\det(A - \lambda I)$ anterior, la multiplicidad algebraica es $n - 1$ para $a - b$ y 1 para $a + (n - 1)b$.

17. $\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = \lambda^2 - (\text{tr } A)\lambda + \det A$. Use la fórmula cuadrática para resolver la ecuación característica:

$$\lambda = \frac{\text{tr } A \pm \sqrt{(\text{tr } A)^2 - 4 \det A}}{2}$$

Los dos valores propios son reales si, y sólo si, el discriminante es no negativo, esto es, $(\text{tr } A)^2 - 4 \det A \geq 0$. Esta desigualdad simplifica a

$$(\text{tr } A)^2 \geq 4 \det A \text{ y } \left(\frac{\text{tr } A}{2}\right)^2 \geq \det A.$$

19. $C_p = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix}$; $\det(C_p - \lambda I) = 6 - 5\lambda + \lambda^2 = p(\lambda)$

21. Si p es un polinomio de segundo orden, entonces un cálculo como el del ejercicio 19 muestra que el polinomio característico de C_p es $p(\lambda) = (-1)^2 p(\lambda)$; así, el resultado es cierto para $n = 2$. Suponga que el resultado es válido para $n = k$ para cierta $k \geq 2$, y considere un polinomio p de grado $k + 1$. Entonces, al desarrollar por cofactores $\det(C_p - \lambda I)$ bajando por la primera columna, el determinante de $C_p - \lambda I$ equivale a

$$(-\lambda)\det\begin{bmatrix} -\lambda & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & & & 1 \\ -a_1 & -a_2 & \cdots & -a_k - \lambda \end{bmatrix} + (-1)^{k+1} a_0$$

La matriz de $k \times k$ que se muestra es $C_q - \lambda I$, donde $q(t) = a_1 + a_2 t + \cdots + a_k t^{k-1} + t^k$. De acuerdo con el supuesto de inducción, el determinante de $C_q - \lambda I$ es $(-1)^k q(\lambda)$. Entonces

$$\begin{aligned}\det(C_p - \lambda I) &= (-1)^{k+1} a_0 + (-\lambda)(-1)^k q(\lambda) \\ &= (-1)^{k+1}[a_0 + \lambda(a_1 + \cdots + a_k \lambda^{k-1} + \lambda^k)] \\ &= (-1)^{k+1} p(\lambda)\end{aligned}$$

Así que la fórmula es válida para $n = k + 1$ cuando es válida para $n = k$. Según el principio de inducción, la fórmula para $\det(C_p - \lambda I)$ es cierta para toda $n \geq 2$.

23. Del ejercicio 22, las columnas de la matriz V de Vandermonde son vectores propios de C_p, correspondientes a los valores propios λ_1, λ_2, λ_3 (las raíces del polinomio p). Puesto que estos valores propios son distintos, los vectores propios forman un conjunto linealmente independiente, según el teorema 2 de la sección 5.1. Por lo tanto, V tiene columnas linealmente independientes y es invertible, de acuerdo con el teorema de la matriz invertible. Por último, dado que las columnas de V son vectores propios de C_p, el teorema de la diagonalización (teorema 5 de la sección 5.3) muestra que $V^{-1} C_p V$ es diagonal.

25. **[M]** Si su programa de matrices calcula valores y vectores propios con métodos iterativos, usted podría encontrar algunas dificultades; aunque $AP - PD$ tenga entradas extremadamente pequeñas y PDP^{-1} sea parecido a A. (Esto era cierto hace algunos años, pero la situación podría cambiar de haber mejorado los programas de matrices.) Si usted estructuró P a partir de los vectores propios del programa, revise el número de condición de P. Éste podría indicar que no se tienen realmente tres vectores propios linealmente independientes.

CAPÍTULO 6

Sección 6.1, página 382

1. 5, 8, $\frac{8}{5}$ **3.** $\begin{bmatrix} 3/35 \\ -1/35 \\ -1/7 \end{bmatrix}$ **5.** $\begin{bmatrix} 8/13 \\ 12/13 \end{bmatrix}$

7. $\sqrt{35}$ **9.** $\begin{bmatrix} -.6 \\ .8 \end{bmatrix}$ **11.** $\begin{bmatrix} 7/\sqrt{69} \\ 2/\sqrt{69} \\ 4/\sqrt{69} \end{bmatrix}$

13. $5\sqrt{5}$ **15.** No ortogonal **17.** Ortogonal

19. Remítase a la *Guía de estudio* (*Study Guide*) después de haber escrito sus respuestas.

21. *Sugerencia*: Use los teoremas 3 y 2 de la sección 2.1.

23. $\mathbf{u}\cdot\mathbf{v} = 0$, $\|\mathbf{u}\|^2 = 30$, $\|\mathbf{v}\|^2 = 101$,
$\|\mathbf{u} + \mathbf{v}\|^2 = (-5)^2 + (-9)^2 + 5^2 = 131 = 30 + 101$

25. El conjunto de todos los múltiplos de $\begin{bmatrix} -b \\ a \end{bmatrix}$ (cuando $\mathbf{v} \neq \mathbf{0}$)

27. *Sugerencia:* Use la definición de ortogonalidad.

29. *Sugerencia:* Considere un vector típico $\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$ en W.

31. *Sugerencia:* Si $\mathbf{x}$ está en $W^\perp$, entonces $\mathbf{x}$ es ortogonal a todo vector en W.

33. **[M]** Formule una conjetura y verifíquela algebraicamente.

Sección 6.2, página 392

1. No ortogonal **3.** No ortogonal **5.** Ortogonal

7. Muestre que $\mathbf{u}_1\cdot\mathbf{u}_2 = 0$, mencione el teorema 4, y observe que dos vectores linealmente independientes en $\mathbb{R}^2$ forman una base. Después obtenga,

$$\mathbf{x} = \tfrac{39}{13}\begin{bmatrix} 2 \\ -3 \end{bmatrix} + \tfrac{26}{52}\begin{bmatrix} 6 \\ 4 \end{bmatrix} = 3\begin{bmatrix} 2 \\ -3 \end{bmatrix} + \tfrac{1}{2}\begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

9. Muestre que $\mathbf{u}_1\cdot\mathbf{u}_2 = 0$, $\mathbf{u}_1\cdot\mathbf{u}_3 = 0$, y $\mathbf{u}_2\cdot\mathbf{u}_3 = 0$. Mencione el teorema 4, y observe que tres vectores linealmente independientes en $\mathbb{R}^3$ forman una base. Después obtenga,

$$\mathbf{x} = \tfrac{5}{2}\mathbf{u}_1 - \tfrac{27}{18}\mathbf{u}_2 + \tfrac{18}{9}\mathbf{u}_3 = \tfrac{5}{2}\mathbf{u}_1 - \tfrac{3}{2}\mathbf{u}_2 + 2\mathbf{u}_3$$

11. $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ **13.** $\mathbf{y} = \begin{bmatrix} -4/5 \\ 7/5 \end{bmatrix} + \begin{bmatrix} 14/5 \\ 8/5 \end{bmatrix}$

15. $\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} .6 \\ -.8 \end{bmatrix}$, la distancia es 1

17. $\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$, $\begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$

19. Ortonormal **21.** Ortonormal

23. Vea la *Guía de estudio* (*Study Guide*).

25. *Sugerencia:* $\|U\mathbf{x}\|^2 = (U\mathbf{x})^T(U\mathbf{x})$. También, los incisos (a) y (c) se deducen de (b).

27. *Sugerencia:* Se necesitan dos teoremas, uno de los cuales sólo es válido para matrices cuadradas.

29. *Sugerencia*: Si se tiene algún candidato para un inverso, puede revisarse que dicho candidato funcione.

31. Suponga que $\hat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$. Reemplace $\mathbf{u}$ por $c\mathbf{u}$ con $c \neq 0$; entonces $\dfrac{\mathbf{y}\cdot(c\mathbf{u})}{(c\mathbf{u})\cdot(c\mathbf{u})}(c\mathbf{u}) = \dfrac{c(\mathbf{y}\cdot\mathbf{u})}{c^2\mathbf{u}\cdot\mathbf{u}}(c)\mathbf{u} = \hat{\mathbf{y}}$

33. Sea $L = \text{Gen}\{\mathbf{u}\}$, donde $\mathbf{u}$ es diferente de cero, y sea $T(\mathbf{x}) = \text{proy}_L\,\mathbf{x}$. Por definición,

$$T(\mathbf{x}) = \frac{\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = (\mathbf{x}\cdot\mathbf{u})(\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u}$$

Para $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$ y cualesquiera escalares c y d, las propiedades del producto interior (teorema 1) muestran que

$$\begin{aligned} T(c\mathbf{x} + d\mathbf{y}) &= [(c\mathbf{x} + d\mathbf{y})\cdot\mathbf{u}](\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u} \\ &= [c(\mathbf{x}\cdot\mathbf{u}) + d(\mathbf{y}\cdot\mathbf{u})](\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u} \\ &= c(\mathbf{x}\cdot\mathbf{u})(\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u} + d(\mathbf{y}\cdot\mathbf{u})(\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u} \\ &= cT(\mathbf{x}) + dT(\mathbf{y}) \end{aligned}$$

Entonces T es lineal.

Sección 6.3, página 400

1. $\mathbf{x} = -\frac{8}{9}\mathbf{u}_1 - \frac{2}{9}\mathbf{u}_2 + \frac{2}{3}\mathbf{u}_3 + 2\mathbf{u}_4$; $\mathbf{x} = \begin{bmatrix} 0 \\ -2 \\ 4 \\ -2 \end{bmatrix} + \begin{bmatrix} 10 \\ -6 \\ -2 \\ 2 \end{bmatrix}$

3. $\begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}$ **5.** $\begin{bmatrix} -1 \\ 2 \\ 6 \end{bmatrix} = \mathbf{y}$

7. $\mathbf{y} = \begin{bmatrix} 10/3 \\ 2/3 \\ 8/3 \end{bmatrix} + \begin{bmatrix} -7/3 \\ 7/3 \\ 7/3 \end{bmatrix}$ **9.** $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 3 \\ -1 \end{bmatrix}$

11. $\begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}$ **13.** $\begin{bmatrix} -1 \\ -3 \\ -2 \\ 3 \end{bmatrix}$ **15.** $\sqrt{40}$

17. a. $U^TU = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $UU^T = \begin{bmatrix} 8/9 & -2/9 & 2/9 \\ -2/9 & 5/9 & 4/9 \\ 2/9 & 4/9 & 5/9 \end{bmatrix}$

b. $\text{proy}_W\,\mathbf{y} = 6\mathbf{u}_1 + 3\mathbf{u}_2 = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$, y $(UU^T)\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$

19. Cualquier múltiplo de $\begin{bmatrix} 0 \\ 2/5 \\ 1/5 \end{bmatrix}$, tal que $\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$

21. Escriba sus respuestas antes de consultar la *Guía de estudio* (*Study Guide*).

23. *Sugerencia:* Use el teorema 3 y el de la descomposición ortogonal. Para la unicidad, suponga que $A\mathbf{p} = \mathbf{b}$ y $A\mathbf{p}_1 = \mathbf{b}$, y considere las ecuaciones $\mathbf{p} = \mathbf{p}_1 + (\mathbf{p} - \mathbf{p}_1)$ y $\mathbf{p} = \mathbf{p} + \mathbf{0}$.

Sección 6.4, página 407

1. $\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix}$ **3.** $\begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3/2 \\ 3/2 \end{bmatrix}$

5. $\begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -4 \\ -1 \end{bmatrix}$ **7.** $\begin{bmatrix} 2/\sqrt{30} \\ -5/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}, \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$

9. $\begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}$ **11.** $\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 3 \\ -3 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 2 \\ -2 \end{bmatrix}$

13. $R = \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix}$

15. $Q = \begin{bmatrix} 1/\sqrt{5} & 1/2 & 1/2 \\ -1/\sqrt{5} & 0 & 0 \\ -1/\sqrt{5} & 1/2 & 1/2 \\ 1/\sqrt{5} & -1/2 & 1/2 \\ 1/\sqrt{5} & 1/2 & -1/2 \end{bmatrix}$, $R = \begin{bmatrix} \sqrt{5} & -\sqrt{5} & 4\sqrt{5} \\ 0 & 6 & -2 \\ 0 & 0 & 4 \end{bmatrix}$

17. Vea la *Guía de estudio* (*Study Guide*).

19. Suponga que $\mathbf{x}$ satisface $R\mathbf{x} = \mathbf{0}$; entonces $QR\mathbf{x} = Q\mathbf{0} = \mathbf{0}$, y $A\mathbf{x} = \mathbf{0}$. Como las columnas de A son linealmente independientes, $\mathbf{x}$ debe ser cero. Este hecho, a su vez, muestra que las columnas de R son linealmente independientes. Puesto que R es cuadrada, resulta ser invertible, según el teorema de la matriz invertible.

21. Denote las columnas de Q mediante $\mathbf{q}_1, \ldots, \mathbf{q}_n$. Observe que $n \leq m$, porque A es $m \times n$ y tiene columnas linealmente independientes. Aplique el hecho de que las columnas de Q pueden ampliarse hasta una base ortonormal para $\mathbb{R}^m$,

por ejemplo, $\{\mathbf{q}_1, \ldots, \mathbf{q}_m\}$. (En la *Guía de Estudio* (*Study Guide*) se describe un método). Sean $Q_0 = [\mathbf{q}_{n+1} \quad \cdots \quad \mathbf{q}_m]$ y $Q_1 = [Q \quad Q_0]$. Entonces, usando la multiplicación de matrices, $Q_1 \begin{bmatrix} R \\ 0 \end{bmatrix} = QR = A$.

23. *Sugerencia:* Parta R como una matriz por bloques de 2×2.

25. **[M]** Las entradas diagonales de R son 20, 6, 10.3923, y 7.0711, hasta cuatro posiciones decimales.

Sección 6.5, página 416

1. a. $\begin{bmatrix} 6 & -11 \\ -11 & 22 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4 \\ 11 \end{bmatrix}$ **b.** $\hat{\mathbf{x}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

3. a. $\begin{bmatrix} 6 & 6 \\ 6 & 42 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$ **b.** $\hat{\mathbf{x}} = \begin{bmatrix} 4/3 \\ -1/3 \end{bmatrix}$

5. $\hat{\mathbf{x}} = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ **7.** $2\sqrt{5}$

9. a. $\hat{\mathbf{b}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ **b.** $\hat{\mathbf{x}} = \begin{bmatrix} 2/7 \\ 1/7 \end{bmatrix}$

11. a. $\hat{\mathbf{b}} = \begin{bmatrix} 3 \\ 1 \\ 4 \\ -1 \end{bmatrix}$ **b.** $\hat{\mathbf{x}} = \begin{bmatrix} 2/3 \\ 0 \\ 1/3 \end{bmatrix}$

13. $A\mathbf{u} = \begin{bmatrix} 11 \\ -11 \\ 11 \end{bmatrix}$, $A\mathbf{v} = \begin{bmatrix} 7 \\ -12 \\ 7 \end{bmatrix}$,

$\mathbf{b} - A\mathbf{u} = \begin{bmatrix} 0 \\ 2 \\ -6 \end{bmatrix}$, $\mathbf{b} - A\mathbf{v} = \begin{bmatrix} 4 \\ 3 \\ -2 \end{bmatrix}$. No, $\mathbf{u}$ no puede ser una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. ¿Por qué?

15. $\hat{\mathbf{x}} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$ **17.** Vea la *Guía de estudio* (*Study Guide*).

19. a. Si $A\mathbf{x} = \mathbf{0}$, entonces $A^TA\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$. Esto demuestra que Nul A está contenido en Nul A^TA.

b. Si $A^TA\mathbf{x} = \mathbf{0}$, entonces $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^T\mathbf{0} = 0$. Así que $(A\mathbf{x})^T(A\mathbf{x}) = 0$, (lo que implica que $\|A\mathbf{x}\|^2 = 0$) y, por lo tanto, $A\mathbf{x} = \mathbf{0}$. Esto demuestra que Nul A^TA está contenido en Nul A.

21. *Sugerencia:* Para (a), aplique un teorema fundamental del capítulo 2.

23. Según el teorema 14, $\hat{\mathbf{b}} = A\hat{\mathbf{x}} = A(A^TA)^{-1}A^T\mathbf{b}$. La matriz $A(A^TA)^{-1}A^T$ se presenta con frecuencia en estadística, donde también se le denomina *matriz-sombrero*.

25. Las ecuaciones normales son $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$, cuya solución es el conjunto de (x, y) tal que $x + y = 3$. Las soluciones corresponden a puntos ubicados sobre la línea a la mitad del camino entre las líneas $x + y = 2$ y $x + y = 4$.

Sección 6.6, página 425

1. $y = .9 + .4x$ **3.** $y = 1.1 + 1.3x$

5. Si dos puntos de datos tienen coordenadas x diferentes, entonces las dos columnas de la matriz de diseño X no pueden ser múltiplos entre sí y, por lo tanto, son linealmente independientes. Según el teorema 14 de la sección 6.5, las ecuaciones normales tienen una solución única.

7. a. $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde $\mathbf{y} = \begin{bmatrix} 1.8 \\ 2.7 \\ 3.4 \\ 3.8 \\ 3.9 \end{bmatrix}$, $X = \begin{bmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 9 \\ 4 & 16 \\ 5 & 25 \end{bmatrix}$,

$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$, $\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}$

b. **[M]** $y = 1.76x - .20x^2$

9. $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde $\mathbf{y} = \begin{bmatrix} 7.9 \\ 5.4 \\ -.9 \end{bmatrix}$, $X = \begin{bmatrix} \cos 1 & \text{sen } 1 \\ \cos 2 & \text{sen } 2 \\ \cos 3 & \text{sen } 3 \end{bmatrix}$,

$\boldsymbol{\beta} = \begin{bmatrix} A \\ B \end{bmatrix}$, $\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}$

11. **[M]** $\beta = 1.45$ y $e = .811$; la órbita es una elipse. La ecuación $r = \beta/(1 - e \cdot \cos\vartheta)$ produce $r = 1.33$ cuando $\vartheta = 4.6$.

13. **[M] a.** $y = -.8558 + 4.7025t + 5.5554t^2 - .0274t^3$

b. La función de velocidad es

$v(t) = 4.7025 + 11.1108t - .0822t^2$, y

$v(4.5) = 53.0$ pies por segundo.

15. *Sugerencia:* Escriba X y $\mathbf{y}$ como en la ecuación (1), y calcule X^TX y $X^T\mathbf{y}$.

17. a. La media de los datos x es $\bar{x} = 5.5$. Los datos, en forma de desviación media, son $(-3.5, 1)$, $(-.5, 2)$, $(1.5, 3)$, $(2.5, 3)$. Las columnas de X son ortogonales porque las entradas de la segunda columna suman 0.

b. $\begin{bmatrix} 4 & 0 \\ 0 & 21 \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 7(5 \end{bmatrix}$,

$y = \frac{9}{4} + \frac{5}{14}x^* = \frac{9}{4} + \frac{5}{14}(x - 5.5)$

19. *Pista:* La ecuación tiene una interpretación geométrica interesante.

Sección 6.7, página 435

1. a. $3, \sqrt{105}, 225$ **b.** Todos los múltiplos de $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$

3. 28 **5.** $5\sqrt{2}, 3\sqrt{3}$ **7.** $\frac{56}{25} + \frac{14}{25}t$

9. a. Polinomio constante, $p(t) = 5$
b. $t^2 - 5$ es ortogonal a p_0 y p_1; valores: $(4, -4, -4, 4)$; respuesta: $q(t) = \frac{1}{4}(t^2 - 5)$

11. $\frac{17}{5}t$

13. Verifique cada uno de los cuatro axiomas. Por ejemplo:

1.	$\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u}) \cdot (A\mathbf{v})$	Definición
	$= (A\mathbf{v}) \cdot (A\mathbf{u})$	Propiedad del producto punto
	$= \langle \mathbf{v}, \mathbf{u} \rangle$	Definición

15.

$\langle \mathbf{u}, c\mathbf{v} \rangle = \langle c\mathbf{v}, \mathbf{u} \rangle$	Axioma 1
$= c\langle \mathbf{v}, \mathbf{u} \rangle$	Axioma 3
$= c\langle \mathbf{u}, \mathbf{v} \rangle$	Axioma 1

17. *Sugerencia:* Calculo cuatro veces el lado derecho.

19. $\langle \mathbf{u}, \mathbf{v} \rangle = \sqrt{a}\sqrt{b} + \sqrt{b}\sqrt{a} = 2\sqrt{ab}$, $\|\mathbf{u}\|^2 = (\sqrt{a})^2 + (\sqrt{b})^2 = a + b$. Puesto que a y b son no negativos, $\|\mathbf{u}\| = \sqrt{a+b}$. De modo semejante, $\|\mathbf{v}\| = \sqrt{b+a}$. De acuerdo con Cauchy-Schwarz, $2\sqrt{ab} \le \sqrt{a+b}\sqrt{b+a} = a + b$. Así, $\sqrt{ab} \le \dfrac{a+b}{2}$.

21. 0 **23.** $2/\sqrt{5}$ **25.** $1, t, 3t^2 - 1$

27. [M] Los nuevos polinomios ortogonales son múltiplos de $-17t + 5t^3$ y $72 - 155t^2 + 35t^4$. Escale estos polinomios para que sus valores en $-2, -1, 0, 1$ y 2 sean enteros pequeños.

Sección 6.8, página 443

1. $y = 2 + \frac{3}{2}t$

3. $p(t) = 4p_0 - .1p_1 - .5p_2 + .2p_3$
$= 4 - .1t - .5(t^2 - 2) + (2\left(\frac{5}{6}t^3 - \frac{17}{6}t\right)$
(Sucede que este polinomio se ajusta exactamente a los datos.)

5. Use la identidad

$$\operatorname{sen} mt \operatorname{sen} nt = \tfrac{1}{2}[\cos(mt - nt) - \cos(mt + nt)]$$

7. Use la identidad $\cos^2 kt = \dfrac{1 + \cos 2kt}{2}$.

9. $\pi + 2 \operatorname{sen} t + \operatorname{sen} 2t + \frac{2}{3} \operatorname{sen} 3t$ [*Sugerencia:* Ahorre tiempo usando resultados del ejemplo 4.]

11. $\frac{1}{2} - \frac{1}{2}\cos 2t$ (¿Por qué?)

13. *Sugerencia:* Tome las funciones f y g en $C[0, 2\pi]$, y fije un entero $m \ge 0$. Escriba el coeficiente de Fourier de $f + g$ que contenga $\cos mt$, y escriba el coeficiente de Fourier que contenga $\operatorname{sen} mt$ $(m > 0)$.

15. [M] La curva cúbica es la gráfica de $g(t) = -.2685 + 3.6095t + 5.8576t^2 - .0477t^3$. La velocidad en $t = 4.5$ segundos es $g'(4.5) = 53.4$ pies por segundo. Esto es, aproximadamente, un 0.7% más rápido que el estimado obtenido en el ejercicio 13 de la sección 6.6.

Capítulo 6 Ejercicios suplementarios, página 444

1. **a.** F **b.** T **c.** T **d.** F **e.** F **f.** T **g.** T **h.** T **i.** F **j.** T **k.** T **l.** F **m.** T **n.** F **o.** F **p.** T **q.** T **r.** F **s.** F

2. *Sugerencia:* Si $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto ortonormal y $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, entonces los vectores $c_1\mathbf{v}_1$ y $c_2\mathbf{v}_2$ son ortogonales, y

$$\|\mathbf{x}\} = \|c_1\mathbf{v}_1 + c_2\mathbf{v}_2\|^2 = \|c_1\mathbf{v}_1\|^2 + \|c_2\mathbf{v}_2\|^2$$
$$= (|c_1|\|\mathbf{v}_1\|)^2 + (|c_2|\|\mathbf{v}_2\|)^2 = |c_1|^2 + |c_2|^2$$

(Explique por qué.) Por lo tanto, la igualdad enunciada es válida para $p = 2$. Suponga que la igualdad es válida para $p = k$, con $k \ge 2$, sea $\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\}$ un conjunto ortonormal, y considere
$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} = \mathbf{u}_k + c_{k+1}\mathbf{v}_{k+1}$, donde $\mathbf{u}_k = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$.

3. Dados $\mathbf{x}$ y un conjunto ortonormal $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ en $\mathbb{R}^n$, sea $\hat{\mathbf{x}}$ la proyección ortogonal de $\mathbf{x}$ sobre el subespacio generado por $\mathbf{v}_1, \dots, \mathbf{v}_p$. De acuerdo con el teorema 10 de la sección 6.3.
$\hat{\mathbf{x}} = (\mathbf{x} \cdot \mathbf{v}_1)\mathbf{v}_1 + \cdots + (\mathbf{x} \cdot \mathbf{v}_p)\mathbf{v}_p$

Según el ejercicio 2, $\|\hat{\mathbf{x}}\|^2 = |\mathbf{x} \cdot \mathbf{v}_1|^2 + \cdots + |\mathbf{x} \cdot \mathbf{v}_p|^2$. La desigualdad de Bessel se deriva del hecho de que $\|\hat{\mathbf{x}}\|^2 \le \|\mathbf{x}\|^2$, lo cual fue señalado antes de la demostración de la desigualdad de Cauchy-Schwarz, en la sección 6.7.

5. Suponga que $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ para todas $\mathbf{x}, \mathbf{y}$ en $\mathbb{R}^n$, y sean $\mathbf{e}_1, \dots, \mathbf{e}_n$ las bases estándar para $\mathbb{R}^n$. Para $j = 1, \dots, n$, $U\mathbf{e}_j$ es la j-ésima columna de U. Como $\|U\mathbf{e}_j\|^2 = (U\mathbf{e}_j) \cdot (U\mathbf{e}_j) = \mathbf{e}_j \cdot \mathbf{e}_j = 1$, las columnas de U son vectores unitarios; puesto que $(U\mathbf{e}_j) \cdot (U\mathbf{e}_k) = \mathbf{e}_j \cdot \mathbf{e}_k = 0$ para $j \ne k$, las columnas son ortogonales por pares.

7. *Sugerencia:* Calcule Q^TQ, usando el hecho de que $(\mathbf{u}\mathbf{u}^T)^T = \mathbf{u}^{TT}\mathbf{u}^T = \mathbf{u}\mathbf{u}^T$.

9. Sea $W = \text{Gen}\{\mathbf{u}, \mathbf{v}\}$. Dado $\mathbf{z}$ en $\mathbb{R}^n$, sea $\hat{\mathbf{z}} = \text{proy}_W \mathbf{z}$. Entonces $\hat{\mathbf{z}}$ está en Col A, donde $A = [\mathbf{u} \quad \mathbf{v}]$, por ejemplo, $\hat{\mathbf{z}} = A\hat{\mathbf{x}}$ para alguna $\hat{\mathbf{x}}$ en $\mathbb{R}^2$. Así que $\hat{\mathbf{x}}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{z}$. Las ecuaciones normales pueden resolverse de modo que produzcan $\hat{\mathbf{x}}$, es posible encontrar entonces $\hat{\mathbf{z}}$ al calcular $A\hat{\mathbf{x}}$.

11. *Sugerencia:* Sean $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}$, y

$A = \begin{bmatrix} \mathbf{v}^T \\ \mathbf{v}^T \\ \mathbf{v}^T \end{bmatrix} = \begin{bmatrix} 1 & -2 & 5 \\ 1 & -2 & 5 \\ 1 & -2 & 5 \end{bmatrix}$. El conjunto de ecuaciones dado es $A\mathbf{x} = \mathbf{b}$, y el conjunto de todas las soluciones por mínimos cuadrados coincide con el conjunto de soluciones de $A^TA\mathbf{x} = A^T\mathbf{b}$ (teorema 13 de la sección 6.5). Estudie esta ecuación y use el hecho de que $(\mathbf{v}\mathbf{v}^T)\mathbf{x} = \mathbf{v}(\mathbf{v}^T\mathbf{x}) = (\mathbf{v}^T\mathbf{x})\mathbf{v}$, porque $\mathbf{v}^T\mathbf{x}$ es un escalar.

13. a. El cálculo fila-columna de $A\mathbf{u}$ muestra que cada fila de A es ortogonal a cada $\mathbf{u}$ en Nul A. Entonces cada fila de A está en $(\text{Nul } A)^\perp$. Como $(\text{Nul } A)^\perp$ es un subespacio, debe contener todas las combinaciones lineales de las filas de A; así que $(\text{Nul } A)^\perp$ contiene Fil A.

b. Si rango $A = r$, entonces dim Nul $A = n - r$, según el teorema del rango. Por el ejercicio 24(c) de la sección 6.3,

$$\dim \text{Nul } A + \dim(\text{Nul } A)^\perp = n$$

Así que $\dim(\text{Nul } A)^\perp$ debe ser r. Pero Fil A es un subespacio de dimensión r de $(\text{Nul } A)^\perp$, según el teorema del rango y el inciso (a). Por lo tanto, Fil A debe coincidir con $(\text{Nul } A)^\perp$.

c. Sustituya A por A^T en el inciso (b) y concluya que Fil A^T coincide con $(\text{Nul } A^T)^\perp$. Como Fil A^T = Col A, esto demuestra (c).

15. Si $A = URU^T$ con U ortogonal, entonces A es semejante a R (porque U es invertible y $U^T = U^{-1}$) y así A tiene los mismos valores propios que R (de acuerdo con el teorema 4 de la sección 5.2), a saber, los n números reales en la diagonal de R.

17. [M] $\dfrac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} = .4618$,

$\text{cond}(A) \times \dfrac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} = 3363 \times (1.548 \times 10^{-4}) = .5206$.

Observe que $\|\Delta\mathbf{x}\|)\,\|\mathbf{x}\|$ casi es igual a $\text{cond}(A)$ veces $\|\Delta\mathbf{b}\|)\,\|\mathbf{b}\|$.

19. [M] $\dfrac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} = 7.178 \times 10^{-8}$, $\dfrac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} = 2.832 \times 10^{-4}$.

Observe que el cambio relativo en $\mathbf{x}$ es *mucho* menor que el cambio relativo en $\mathbf{b}$. De hecho, como

$$\text{cond}(A) \times \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} = 23683 \times (2.832 \times 10^{-4}) = 6.707$$

la cota teórica para los cambios relativos en $\mathbf{x}$ es de 6.707 (hasta cuatro cifras significativas). Este ejercicio muestra que aún cuando un número de condición es grande, el error relativo en una solución no tiene que ser tan grande como pudiera esperarse.

CAPÍTULO 7

Sección 7.1, página 454

3. No simétrica **5.** No simétrica

7. Ortogonal, $\begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}$ **9.** No ortogonal

11. Ortogonal, $\begin{bmatrix} 2/3 & 0 & 5/\sqrt{45} \\ 2/3 & 1/\sqrt{5} & -4/\sqrt{45} \\ 1/3 & -2/\sqrt{5} & -2/\sqrt{45} \end{bmatrix}$

13. $P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$, $D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$

15. $P = \begin{bmatrix} -4/\sqrt{17} & 1/\sqrt{17} \\ 1/\sqrt{17} & 4/\sqrt{17} \end{bmatrix}$, $D = \begin{bmatrix} 17 & 0 \\ 0 & 0 \end{bmatrix}$

17. $P = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix}$, $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

19. $P = \begin{bmatrix} -1/\sqrt{5} & 4/\sqrt{45} & -2/3 \\ 2/\sqrt{5} & 2)\sqrt{45} & -1/3 \\ 0 & 5)\sqrt{45} & 2/3 \end{bmatrix}$,

$D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

21. $P = \begin{bmatrix} .5 & -.5 & -1/\sqrt{2} & 0 \\ .5 & .5 & 0 & -1/\sqrt{2} \\ .5 & -.5 & 1/\sqrt{2} & 0 \\ .5 & .5 & 0 & 1/\sqrt{2} \end{bmatrix}$,

$D = \begin{bmatrix} 9 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

23. $P = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix}$, $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

25. Vea la *Guía de estudio* (*Study Guide*).

27. $(B^TAB)^T = B^TA^TB^{TT}$ Producto de transpuestas en orden inverso

$= B^TAB$ Porque A es simétrica

El resultado en torno a B^TB es un caso especial cuando $A = I$. $(BB^T)^T = B^{TT}B^T = BB^T$, así que BB^T es simétrica.

29. *Sugerencia:* Use una diagonalización ortogonal de A, o recurra al teorema 2.

31. El teorema de la diagonalización presentado en la sección 5.3 postula que las columnas de P son vectores propios (linealmente independientes) correspondientes a los valores propios de A enlistados en la diagonal de D. Así que P tiene exactamente k columnas de vectores propios correspondientes a λ. Estas k columnas forman una base para el espacio propio.

33. $A = 8\mathbf{u}_1\mathbf{u}_1^T + 6\mathbf{u}_2\mathbf{u}_2^T + 3\mathbf{u}_3\mathbf{u}_3^T$

$$= 8\begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 6\begin{bmatrix} 1/6 & 1/6 & -2/6 \\ 1/6 & 1/6 & -2/6 \\ -2/6 & -2/6 & 4/6 \end{bmatrix} + 3\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

35. *Sugerencia:* $(\mathbf{u}\mathbf{u}^T)\mathbf{x} = \mathbf{u}(\mathbf{u}^T\mathbf{x}) = (\mathbf{u}^T\mathbf{x})\mathbf{u}$, porque $\mathbf{u}^T\mathbf{x}$ es un escalar.

Sección 7.2, página 462

1. a. $5x_1^2 + \frac{2}{3}x_1x_2 + x_2^2$ **b.** 185 **c.** 16

3. a. $\begin{bmatrix} 10 & -3 \\ -3 & -3 \end{bmatrix}$ **b.** $\begin{bmatrix} 5 & 3/2 \\ 3/2 & 0 \end{bmatrix}$

5. a. $\begin{bmatrix} 8 & -3 & 2 \\ -3 & 7 & -1 \\ 2 & -1 & -3 \end{bmatrix}$ **b.** $\begin{bmatrix} 0 & 2 & 3 \\ 2 & 0 & -4 \\ 3 & -4 & 0 \end{bmatrix}$

7. $\mathbf{x} = P\mathbf{y}$, donde $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\mathbf{y}^T D\mathbf{y} = 6y_1^2 - 4y_2^2$

En los ejercicios 9 a 14, son posibles otras respuestas (cambio de variables y nueva forma cuadrática).

9. Definida positiva; los valores propios son 7 y 2
Cambio de variable: $\mathbf{x} = P\mathbf{y}$, con $P = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$
Nueva forma cuadrática: $7y_1^2 + 2y_2^2$

11. Indefinida; los valores propios son 7 y -3
Cambio de variable: $\mathbf{x} = P\mathbf{y}$, con $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$
Nueva forma cuadrática: $7y_1^2 - 3y_2^2$

13. Semidefinida positiva; los valores propios son: 10 y 0
Cambio de variable: $\mathbf{x} = P\mathbf{y}$, con $P = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}$
Nueva forma cuadrática: $10y_1^2$

15. **[M]** Semidefinida negativa; los valores propios son 0, -6, -8, -12 Cambio de variable: $\mathbf{x} = P\mathbf{y}$;

$$P = \begin{bmatrix} 3/\sqrt{12} & 0 & -1/2 & 0 \\ 1/\sqrt{12} & -2/\sqrt{6} & 1/2 & 0 \\ 1/\sqrt{12} & 1/\sqrt{6} & 1/2 & -1/\sqrt{2} \\ 1/\sqrt{12} & 1/\sqrt{6} & 1/2 & 1/\sqrt{2} \end{bmatrix}$$

Nueva forma cuadrática: $-6y_2^2 - 8y_3^2 - 12y_4^2$

17. **[M]** Indefinida; los valores propios son 8.5 y -6.5
Cambio de variable: $\mathbf{x} = P\mathbf{y}$;

$$P = \frac{1}{\sqrt{50}}\begin{bmatrix} 3 & -4 & 3 & 4 \\ 5 & 0 & -5 & 0 \\ 4 & 3 & 4 & -3 \\ 0 & 5 & 0 & 5 \end{bmatrix}$$

Nueva forma cuadrática: $8.5y_1^2 + 8.5y_2^2 - 6.5y_3^2 - 6.5y_4^2$

19. 8 **21.** Vea la *Guía de estudio* (*Study Guide*).

23. Escriba el polinomio característico de dos maneras:

$$\det(A - \lambda I) = \det\begin{bmatrix} a-\lambda & b \\ b & d-\lambda \end{bmatrix} = \lambda^2 - (a+d)\lambda + ad - b^2$$

y

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2$$

Iguale los coeficientes para obtener $\lambda_1 + \lambda_2 = a + d$ y $\lambda_1\lambda_2 = ad - b^2 = \det A$.

25. El ejercicio 27 de la sección 7.1 mostró que B^TB es simétrica. También, $\mathbf{x}^TB^TB\mathbf{x} = (B\mathbf{x})^TB\mathbf{x} = \|B\mathbf{x}\|^2 \geq 0$, entonces la forma cuadrática es semidefinida positiva, y se afirma que la matriz B^TB es semidefinida positiva. *Sugerencia:* Para mostrar que B^TB es definida positiva cuando B es cuadrada e invertible, suponga que $\mathbf{x}^TB^TB\mathbf{x} = 0$ y deduzca que $\mathbf{x} = \mathbf{0}$.

27. *Sugerencia:* Demuestre que $A + B$ es simétrica y que la forma cuadrática $\mathbf{x}^T(A + B)\mathbf{x}$ es definida positiva.

Sección 7.3, página 470

1. $\mathbf{x} = P\mathbf{y}$, donde $P = \begin{bmatrix} 1/3 & 2/3 & -2/3 \\ 2/3 & 1/3 & 2/3 \\ -2/3 & 2/3 & 1/3 \end{bmatrix}$

3. a. 9 **b.** $\pm\begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix}$ **c.** 6

5. a. 7 **b.** $\pm\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ **c.** 3

7. $\pm\begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$ **9.** $5 + \sqrt{5}$ **11.** 3

13. *Sugerencia:* Si $m = M$, tome $\alpha = 0$ en la fórmula de $\mathbf{x}$. Esto es, sea $\mathbf{x} = \mathbf{u}_n$, y verifique si $\mathbf{x}^T A\mathbf{x} = m$. Si $m < M$ y si t es un número entre m y M, entonces $0 \le t - m \le M - m$ y $0 \le (t - m)/(M - m) \le 1$. Así, sea $\alpha = (t - m)/(M - m)$. Resuelva la expresión para α para ver que $t = (1 - \alpha)m + \alpha M$. Mientras α va de 0 a 1, t va de m a M. Estructure $\mathbf{x}$ como en el enunciado del ejercicio y verifique sus propiedades.

15. [M] **a.** 7.5 **b.** $\begin{bmatrix} .5 \\ .5 \\ .5 \\ .5 \end{bmatrix}$ **c.** $-.5$

17. [M] **a.** -4 **b.** $\begin{bmatrix} -3/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \end{bmatrix}$ **c.** -10

Sección 7.4, página 481

1. 3, 1 **3.** 3, 2

Las respuestas de los ejercicios 5 a 13 no son las únicas posibilidades.

5. $\begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$

9. $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}$
$\times \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$

11. $\begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$
$\times \begin{bmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix}$

13. $\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$
$= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$
$\times \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{18} & 1/\sqrt{18} & -4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{bmatrix}$

15. **a.** rango $A = 2$

b. Base para Col A: $\begin{bmatrix} .40 \\ .37 \\ -.84 \end{bmatrix}, \begin{bmatrix} -.78 \\ -.33 \\ -.52 \end{bmatrix}$

Base para Nul A: $\begin{bmatrix} .58 \\ -.58 \\ .58 \end{bmatrix}$

(Recuerde que V^T aparece en la DVS.)

17. Sea $A = U\Sigma V^T = U\Sigma V^{-1}$. Puesto que A es cuadrada e invertible, rango $A = n$ y todas las entradas en la diagonal de Σ deben ser diferentes a cero. Así que $A^{-1} = (U\Sigma V^{-1})^{-1} = V\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$.

19. *Sugerencia:* Como U y V son ortogonales

$$A^TA = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V(\Sigma^T\Sigma)V^{-1}$$

Así que V diagonaliza A^TA. ¿Qué plantea esto acerca de V?

21. Sea $A = U\Sigma V^T$. La matriz PU es ortogonal, porque tanto P como U son ortogonales. (Vea el ejercicio 29 de la sección 6.2.) Por lo tanto, la ecuación $PA = (PU)\Sigma V^T$ tiene la forma requerida para una descomposición en valores singulares. Según el ejercicio 19, las entradas diagonales en Σ son los valores singulares de PA.

23. *Sugerencia:* Use un desarrollo columna-fila de $(U\Sigma)V^T$.

25. *Sugerencia:* Considere la DVS de la matriz estándar de T —por ejemplo, $A = U\Sigma V^T = U\Sigma V^{-1}$. Sean $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ y $\mathcal{C} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ bases estructuradas a partir de las columnas de V y U, respectivamente—. Encuentre la matriz para T relativa a $\mathcal{B}$ y $\mathcal{C}$, como en la sección 5.4. Para hacer esto, debe demostrarse que $V^{-1}\mathbf{v}_j = \mathbf{e}_j$, la j-ésima columna de I_n.

27. [M] $\begin{bmatrix} -.57 & -.65 & -.42 & .27 \\ .63 & -.24 & -.68 & -.29 \\ .07 & -.63 & .53 & -.56 \\ -.51 & .34 & -.29 & -.73 \end{bmatrix}$
$\times \begin{bmatrix} 16.46 & 0 & 0 & 0 & 0 \\ 0 & 12.16 & 0 & 0 & 0 \\ 0 & 0 & 4.87 & 0 & 0 \\ 0 & 0 & 0 & 4.31 & 0 \end{bmatrix}$
$\times \begin{bmatrix} -.10 & .61 & -.21 & -.52 & .55 \\ -.39 & .29 & .84 & -.14 & -.19 \\ -.74 & -.27 & -.07 & .38 & .49 \\ .41 & -.50 & .45 & -.23 & .58 \\ -.36 & -.48 & -.19 & -.72 & -.29 \end{bmatrix}$

29. [M] 25.9343, 16.7554, 11.2917, 1.0785, .0037793; $\sigma_1/\sigma_5 = 68{,}622$

Sección 7.5, página 489

1. $M = \begin{bmatrix} 12 \\ 10 \end{bmatrix}$; $B = \begin{bmatrix} 7 & 10 & -6 & -9 & -10 & 8 \\ 2 & -4 & -1 & 5 & 3 & -5 \end{bmatrix}$;

$S = \begin{bmatrix} 86 & -27 \\ -27 & 16 \end{bmatrix}$

3. $\begin{bmatrix} .95 \\ -.32 \end{bmatrix}$ para $\lambda = 95.2$, $\begin{bmatrix} .32 \\ .95 \end{bmatrix}$ para $\lambda = 6.8$

5. [M] (.130, .874, .468), el 75.9% de la varianza.

7. $y_1 = .95x_1 - .32x_2$; y_1 explica el 93.3% de la varianza.

9. $c_1 = 1/3$, $c_2 = 2/3$, $c_3 = 2/3$; la varianza de y es 9.

11. a. Si $\mathbf{w}$ es el vector en $\mathbb{R}^N$ con un 1 en cada posición, entonces

$$[\,\mathbf{X}_1 \quad \cdots \quad \mathbf{X}_N\,]\,\mathbf{w} = \mathbf{X}_1 + \cdots + \mathbf{X}_N = \mathbf{0} \text{ m}$$

porque las $\mathbf{X}_k$ están en forma de desviación media. Entonces

$$\begin{aligned} &[\,\mathbf{Y}_1 \quad \cdots \quad \mathbf{Y}_N\,]\,\mathbf{w} \\ &= [\,P^T\mathbf{X}_1 \quad \cdots \quad P^T\mathbf{X}_N\,]\,\mathbf{w} \qquad \text{Por definición} \\ &= P^T[\,\mathbf{X}_1 \quad \cdots \quad \mathbf{X}_N\,]\,\mathbf{w} = P^T\mathbf{0} = \mathbf{0} \end{aligned}$$

Esto es, $\mathbf{Y}_1 + \cdots + \mathbf{Y}_N = \mathbf{0}$, así que las $\mathbf{Y}_k$ están en forma de desviación media.

b. *Sugerencia*: Como las $\mathbf{X}_j$ están en forma de desviación media, la matriz de covarianza de las $\mathbf{X}_j$ es

$$1/(N-1)\,[\,\mathbf{X}_1 \quad \cdots \quad \mathbf{X}_N\,]\,[\,\mathbf{X}_1 \quad \cdots \quad \mathbf{X}_N\,]^T$$

Encuentre la matriz de covarianza de las $\mathbf{Y}_j$, usando el inciso (a).

13. Si $B = [\,\hat{\mathbf{X}}_1 \quad \cdots \quad \hat{\mathbf{X}}_N\,]$, entonces

$$\begin{aligned} S &= \frac{1}{N-1}BB^T = \frac{1}{N-1}[\,\hat{\mathbf{X}}_1 \quad \cdots \quad \hat{\mathbf{X}}_n\,]\begin{bmatrix} \hat{\mathbf{X}}_1^T \\ \vdots \\ \hat{\mathbf{X}}_N^T \end{bmatrix} \\ &= \frac{1}{N-1}\sum_1^N \hat{\mathbf{X}}_k\hat{\mathbf{X}}_k^T = \frac{1}{N-1}\sum_1^N (\mathbf{X}_k - \mathbf{M})(\mathbf{X}_k - \mathbf{M})^T \end{aligned}$$

Capítulo 7 Ejercicios suplementarios, página 491

1. **a.** T **b.** F **c.** T **d.** F **e.** F
f. F **g.** F **h.** T **i.** F **j.** F
k. F **l.** F **m.** T **n.** F **o.** T
p. T **q.** F

3. Si rango $A = r$, entonces dim Nul $A = n - r$, según el teorema del rango. Entonces 0 es un valor propio con multiplicidad $n - r$. Por lo tanto, de los n términos de la descomposición espectral de A, exactamente $n - r$ son cero. Los r términos restantes (correspondientes a los valores propios diferentes de cero) son todos matrices con rango 1, tal como se mencionó en el análisis de la descomposición espectral.

5. Si $A\mathbf{v} = \lambda\mathbf{v}$ para alguna λ diferente de cero, entonces $\mathbf{v} = \lambda^{-1}A\mathbf{v} = A(\lambda^{-1}\mathbf{v})$, lo cual muestra que $\mathbf{v}$ es una combinación lineal de las columnas de A.

7. *Sugerencia:* Si $A = R^TR$, donde R es invertible, entonces A es definida positiva, según el ejercicio 25 de la sección 7.2. A la inversa, suponga que A es definida positiva. Entonces, de acuerdo con el ejercicio 26 de la sección 7.2, $A = B^TB$ para alguna matriz definida positiva B. Explique por qué B admite una factorización QR, y use ésta para crear la factorización Cholesky de A.

9. Si A es de $m \times n$ y $\mathbf{x}$ está en $\mathbb{R}^n$, entonces $\mathbf{x}^TA^TA\mathbf{x} = (A\mathbf{x})^T(A\mathbf{x}) = \|A\mathbf{x}\|^2 \geq 0$. Por lo tanto, A^TA es semidefinida positiva. Según el ejercicio 22 de la sección 6.5, rango A^TA = rango A.

11. *Sugerencia:* Escriba una DVS de A en la forma $A = U\Sigma V^T = PQ$, donde $P = U\Sigma U^T$ y $Q = UV^T$. Muestre que P es simétrica y que tiene los mismos valores propios que Σ. Explique por qué es Q una matriz ortogonal.

13. a. Si $\mathbf{b} = A\mathbf{x}$, entonces $\mathbf{x}^+ = A^+\mathbf{b} = A^+A\mathbf{x}$. Según el ejercicio 12(a), $\mathbf{x}^+$ es la proyección ortogonal de $\mathbf{x}$ sobre Fil A.

b. De (a) y del ejercicio 12(c),

$$A\mathbf{x}^+ = A(A^+A\mathbf{x}) = (AA^+A)\mathbf{x} = A\mathbf{x} = \mathbf{b}.$$

c. Como $\mathbf{x}^+$ es la proyección ortogonal sobre Fil A, el teorema de Pitágoras muestra que $\|\mathbf{u}\|^2 = \|\mathbf{x}^+\|^2 + \|\mathbf{u} - \mathbf{x}^+\|^2$. El inciso (c) se deduce inmediatamente.

15. [M] $A^+ = \dfrac{1}{40}\cdot\begin{bmatrix} -2 & -14 & 13 & 13 \\ -2 & -14 & 13 & 13 \\ -2 & 6 & -7 & -7 \\ 2 & -6 & 7 & 7 \\ 4 & -12 & -6 & -6 \end{bmatrix}$, $\hat{\mathbf{x}} = \begin{bmatrix} .7 \\ .7 \\ -.8 \\ .8 \\ .6 \end{bmatrix}$

La forma escalonada reducida de $\begin{bmatrix} A \\ \mathbf{x}^T \end{bmatrix}$ es igual a la forma escalonada reducida de A, excepto por una fila extra de ceros. Así que sumar múltiplos escalares de las filas de A a $\mathbf{x}^T$ puede producir el vector cero, lo cual demuestra que $\mathbf{x}^T$ está en Fil A.

Bases para Nul A: $\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$

Índice

¿NO OLVIDAS ALGO?

Al comprar este libro de texto, **Pearson Educación** te da acceso a la tecnología más avanzada para complementar tu aprendizaje, dentro y fuera del salón de clases.

Acompañando a este libro, puedes encontrar cuestionarios de autoevaluación, ejercicios interactivos, animaciones, casos de estudio, resúmenes o hasta un curso en línea dentro de nuestra plataforma ***CourseCompass****.

Consulta la página Web del libro para conocer los recursos que están disponibles. O pregunta a tu profesor sobre el material que puso a tu disposición para el curso y **entrégale el formulario que está al reverso para solicitar tu código de acceso**.

¡No dejes pasar esta oportunidad y únete a los millones de alumnos que están sacando el máximo provecho de su libro de texto!

*****CourseCompass*** es una plataforma educativa en línea desarrollada por Blackboard Technologies® exclusivamente para **Pearson Educación**.

SOLICITUD DE CÓDIGO DE ACCESO PARA *COURSECOMPASS*

DATOS DEL ALUMNO

Nombre completo | e-mail

DATOS DE LA INSTITUCIÓN

Nombre de la institución | Campus o Facultad

Dirección | Ciudad y estado | País

Nombre del profesor | e-mail del profesor

Nombre de la materia | Grado (Nº semestre, trimestre, etc.) | Nombre de la carrera

DATOS DEL LIBRO

Título | Edición | Autor | ISBN

¿Es el texto principal? Sí ○ No ○

¿Dónde adquiriste el libro?

¿Consideras adecuado el precio? Sí ○ No ○

¿Cuentas con una computadora propia? Sí ○ No ○

¿Cuentas con acceso a Internet? Sí ○ No ○

¿Cuentas con laboratorio de cómputo en tu escuela? Sí ○ No ○

¿Has utilizado anteriormente esta u otra plataforma en línea? Sí ○ No ○

¿Cuál?

¿Ayudó a mejorar tu desempeño? Sí ○ No ○

¿Por qué?

¿Utilizas actualmente algún otro libro de Pearson Educación? Sí ○ No ○

¿Cuáles?

1. Título | edición | Autor

Materia | Profesor | ¿CourseCompass? Sí ○ No ○

2. Título | edición | Autor

Materia | Profesor | ¿CourseCompass? Sí ○ No ○

PARA LLENAR POR EL PROFESOR

(Llenar una sola por grupo y entregar al frente con el resto de las solicitudes)

Clave del curso (Course ID)[1] | ISBN del curso[1]

Fecha de inicio del curso / / | Culminación / / | Límite para registro de alumnos[2] / /

Número de códigos solicitados | Total de alumnos en el grupo | Teléfono de contacto

¿Existe el libro en biblioteca? Sí ○ No ○ | Fecha de entrega de solicitudes / /

¿Le gustaría recibir información sobre otros materiales de Pearson Educación? Sí ○ No ○

[1] Entre a la sección **Course List** haciendo clic en la pestaña **Courses** de CourseCompass. La información aparece debajo de cada curso de su lista.

[2] En la sección **Course List**, hacer clic en el botón **Course Settings** de este curso y luego en la liga **Course Dates**. La fecha límite para inscripción aparece como **Enrollment End Date**.

Para mayor información, entre a **www.pearsoneducacion.net/coursecompass**
o escríbanos a **editorialmx@pearsoned.com**